

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 14, 2005, 17:06:19 ; Search time 168.5 Seconds  
(without alignments)  
4448.318 Million cell updates/sec

Title: US-09-905-083A-30  
Perfect score: 1780

Sequence: 1 ggatttcgggctccatggc.....aagaacacacaaacccctcag 969

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame+\_n2p.model -DEV=slp  
-Q=/cgm2\_1/USPTO\_spool\_p/US9905083/runat\_14072005\_140352\_10602/app\_query.fasta\_1.1159  
-DB=A Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9905083@cgm2\_1\_224@runat\_14072005\_140352\_10602 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

A Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1364	76.6	253	2 AAR67888	Aar67888 Human str
2	1364	76.6	253	2 AAW05383	Aaw05383 Human amy
3	1364	76.6	253	5 ABB84421	Abb84421 Human SCC
4	1364	76.6	253	5 ABB84406	Abb84406 Human SCC
5	1364	76.6	253	5 AAU82740	Aau82740 Amino aci
6	1364	76.6	253	6 ABU07440	Abu07440 Protein d
7	1364	76.6	253	6 ABU07471	Abu07471 Protein d
8	1364	76.6	253	6 ABR58471	Abr58471 Human str
9	1364	76.6	253	7 ADB80484	Adb80484 Ovarian c
10	1364	76.6	253	7 ADJ68833	Adj68833 Human hea

11	1364	76.6	253	7 ADN39180	Adn39180 Cancer/an
12	1364	76.6	253	8 ADL06515	Adl06515 Human tum
13	1364	76.6	253	8 ADN04182	Adn04182 Antipsoori
14	1364	76.6	253	8 ADR72880	Adr72880 Human ova
15	1350	75.8	250	6 ADA05732	Ada05732 Human NOV
16	1344	75.5	250	8 ADN62896	Adn62896 Human NOV
17	1337	75.1	257	3 AAB21326	Aab21326 Human HSC
18	1252	70.3	247	6 ADA05742	Ada05742 Human NOV
19	1252	70.3	247	8 ADN62906	Adn62906 Human NOV
20	1252	70.3	252	8 ADA05734	Ada05734 Human NOV
21	1252	70.3	252	8 ADN62898	Adn62898 Human NOV
22	1235	69.4	225	4 AAB98502	Aab98502 Human Str
23	1119	62.9	224	6 ADA05744	Ada05744 Human NOV
24	1119	62.9	224	8 ADN62908	Adn62908 Human NOV
25	1085.5	61.0	249	5 ABB84420	Abb84420 Porcine S
26	1035.5	58.2	198	8 ADA05736	Ada05736 Human NOV
27	1035.5	58.2	198	8 ADN62900	Adn62900 Human NOV
28	1011.5	56.8	249	5 ABB84423	Abb84423 Murine SC
29	996.5	56.0	243	5 ABB84419	Abb84419 Bovine SC
30	940.5	52.8	226	5 ABB84422	Abb84422 Rat SCCE
31	937.5	52.7	181	6 ADA05738	Ada05738 Human NOV
32	937.5	52.7	181	8 ADN62902	Adn62902 Human NOV
33	773	43.4	144	8 ADI39727	Adi39727 Stratum c
34	773	43.4	144	8 ADI37151	Adi37151 Stratum c
35	641.5	36.0	136	4 ABG23378	Abg23378 Novel hum
36	614	34.5	260	2 AAW10594	Aaw10594 Human rec
37	614	34.5	260	2 AAW12393	Aaw12393 Mouse neu
38	614	34.5	260	5 ABB57219	Abb57219 Mouse lsc
39	614	34.5	260	5 ADI17073	Adi17073 Murine NO
40	614	34.5	260	8 ADI39731	Adi39731 Mouse neu
41	614	34.5	260	8 ADI37155	Adi37155 Mouse neu
42	612.5	34.4	293	2 AAY16777	Aay16777 Human ker
43	612.5	34.4	293	2 AAY30524	Aay30524 Human sec
44	612.5	34.4	293	2 AAY38412	Aay38412 Human sec
45	612.5	34.4	293	2 AAY38426	Aay38426 Human sec

## ALIGNMENTS

### RESULT 1

AAR67888

ID AAR67888 standard; protein; 253 AA.

AC AAR67888;

DT 25-MAR-2003 (revised)

DT 09-AUG-1995 (first entry)

DS Human stratum corneum chymotrophic recombinant enzyme (SCCE).

KW Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
KW callosities; keratosis pilaris; ichthyoses; eczema.

OS Homo sapiens.

PN WO9500651-A1.

PD 05-JAN-1995.

PF 20-JUN-1994; 94WO-IB000166.

PR 18-JUN-1993; 93DK-00000725.

PA (SYMB-) SYMBICOM AB.

PI Egelrud T, Hansson L;

DR WPI; 1995-052088/07.

DR N-PSDB; AAQ81203.

XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
XX related vectors, transformed cells and polypeptides, useful for treating  
XX skin disorders, e.g. acne or psoriasis, and for identification of

PT specific inhibitors.

XX Disclosure; Page 97; 137pp; English.

XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
CC and skin care products, especially to treat and prevent acne, xeroderma,  
CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
CC recombinantly following mammal, insect, plant, or microorganism  
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
CC field.)

XX Sequence 253 AA;

Alignment Scores:  
Pred. No.: 1,02e-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 2 Gaps: 0

US-09-905-083A-30 (1-969) x AAR67888 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTGCCCCCTGCAGATCTTACTGTATCTTACTGTCTTGGAACT 75  
DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGGTGACAGATTATTGATGGCGGCCCATGTGCAAGAGGTCC 135  
DB 21 AlaGlyGluGluAlaGlnGlyAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCCTGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 195  
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCGCTGGTGCTCACTGCGCCCACTGCAAGATGAATGATGATGATGATGATGAT 255  
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluThrValHisLeu 80  
QY 256 GCGAGTATAGCTGGGCGACAGGAGCTCAGAGGATCAAGGCTCGAATCTTTCGCG 315  
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCCATGTTAATGACCTCATGCTGCTGCTGCTGCTGCTGCT 375  
DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGCCAGGCTGTTCATCCATGCTGAGAAAGTCAAGGCTGCGCTCCCGCTGCGAAACCCCT 435  
DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCGGCTGGGGGCTACTACAGGAGCCAGATGTGACCTTCCC 495  
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGGTGATGTCAGCTCATCTCCCGCCAGGACTGCACCAAGGTTTAC 555  
DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTATCGAAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615  
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200  
QY 616 TGCATGTGTACTCAGGGGGACCGTTGTTGTGTCAGAGGTACCTGCAAGTCTGGTGCTCC 675  
DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGGAATCTTCCCTTGGGGCAACCCCAATGACCCAGGAGTCTACCTCAAGTGTGCAAG 735  
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAAAAGCATCGC 774

DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 2

AAW05383

ID AAW05383 standard; protein; 253 AA.

XX AC AAW05383;

XX DT 31-DEC-1996 (first entry)

XX DE Human amyloid precursor protein protease.

XX KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;

XX OS Homo sapiens.

XX PN WO9631122-A1.

XX PD 10-OCT-1996.

XX PF 02-APR-1996; 96WO-US004294.

XX PR 04-APR-1995; 95US-00416257.

XX PA (ELIL) LILLY & CO ELI.

XX PI Dixon BP, Johnstone EM, Little SP;

XX DR WPI; 1996-464694/46.

XX DR N-PSDB; AAT39783.

XX PT New isolated human amyloid precursor protein protease - used to develop  
prods. for the treatment or diagnosis of associated conditions, esp.  
Alzheimer's disease.

XX PS Claim 1; Page 44-45; 55pp; English.

XX CC Human amyloid precursor protein protease (AAW05383) is involved in the  
processing or clearance of amyloid precursor protein to form beta-amyloid  
peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
obtd. from a human lung library. Recombinant protease can be produced in  
transfected or transfected prokaryotic (partic. E. coli) or eukaryotic  
(partic. AV-120 host cells. It is used to develop products for the design  
and testing of cpds. useful for treating or preventing conditions  
associated with beta-amyloid peptide, esp. Alzheimer's disease

XX SQ Sequence 253 AA;

Alignment Scores:

Pred. No.: 1,02e-134 Length: 253

Score: 1364.00 Matches: 253

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 76.63% Indels: 0

DB: 2 Gaps: 0

US-09-905-083A-30 (1-969) x AAW05383 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTGCCCCCTGCAGATCTTACTGTATCTTGGAACT 75

DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20

QY 76 GCAGGAGAGAGCCAGGGTGACAGATTATTGATGGCGGCCCATGTGCAAGAGGTCC 135

DB 21 AlaGlyGluGluAlaGlnGlyAspGlyAlaProCysAlaArgGlySer 40

QY 136 CACCCTGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 195

DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

QY 196 AATGAGCGCTGGTGCTTCACTGCGCGCCCACTGCAAGATGAATGATGATGATGATGATGAT 255



Qy	496	TCTGACCTCATGTGCTGGATGTC	CAAGTCATCTCCCCCCAGGACTCC	CAAGAAGGTTTAC	555
Db	161	SerAspLeuMetCysValAspVal	LeuIleSerProGlnAspCys	ThrIysVal	180
Qy	556	AAGGACTTACTGCGAAATTC	ATGCTGTGCTGGCATCCCCGACT	CAAGAAAAACGCC	615
Db	181	LysAspLeuLeuGluAsnSerMet	LeuCysAlaGlyIleProAsp	SerIysLysAsnAla	200
Qy	616	TGCAATGTGTACTCAGGGGG	CCGTTGTGTGCAGAGGTAC	CTCTGCAAGGTCTGGTGTC	675
Db	201	CysAsnGlyAspSerGlyGly	ProLeuValCysArgGlyThr	LeuGlnGlyLeuValSer	220
Qy	676	TGGGNACTTTCCCTTGGGG	CCACCCTAATGACCCAGGAGT	CTACACTCAAGTGTCCAG	735
Db	221	TrpGlyThrPheProCysGly	GlnProAsnAspProGly	ValIyThrGlnValCysLys	240
Qy	736	TTCCACCAAGTGGATAAAT	GCACACCATGAAAAAGCATCGC		774
Db	241	PheThrLysTrpIleAsnAsp	ThrMetLysLysHisArg		253

abnormal or unwanted phenotype, and for screening or identifying a compound or composition effective for the prevention or treatment of inflammatory skin diseases selected from diseases consisting of epidermal hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation, pruritus, atopic dermatitis, eczema, acne and inherited skin diseases with epidermal hyperkeratosis. The mammal of the invention is also useful as a model for further studies of itch mechanisms and the testing of potential compounds and compositions for relieve of various skin diseases where itch is a component. This sequence represents the human stratum corneum chymotryptic enzyme, SCCB which is a serine protease synonymous with human kallikrein 7 (KUK7) and is used in the development of the transgenic mammals described in the invention

Sequence 253 AA;

Alignment Scores:	
Pred. No.:	1.02e-134
Score:	1364.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	76.63%
DB:	5
	0
Length:	253
Matches:	253
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-905-083A-30 (1-969) x ABB84406 (1-253)

Qy	16	ATGCAGATCCCTTCCTCGCCCTCGAGATCTTACGTGCTATCTCTTACGCTTGAAACT	75
Db	1	MetAlaargSerLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr	20
Qy	76	GCAGAGAAAGCCAGGCTGACAAGATTATTGATGGCGCCCTCATGTGCAAGAGCTCC	135
Db	21	AlaGlyGluAlaGlnGlyAspIleIleAspGlyAlaProCysAlaargGlySer	40
Qy	136	CACCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGGCTCTGGTGC	195
Db	41	HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal	60
Qy	196	AATCAGCGCTGGGTGCTCACTCGCCGCCACATGCAAGATGAATGACTACACCGTGACACCTG	255
Db	61	AsnGluargTrpValLeuThrAlaAlaHisCysIysMetAsnGluTrpThrValHisIleu	80
Qy	256	GGCAGTGATACGCTGGGCGACAGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG	315
Db	81	GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg	100
Qy	316	CACCCCGCTACTCCACACAGACCATGTTAATGACCTCAGCTCAGCTCGTGAAGCTCAATAGC	375
Db	101	HisProGlyTrpSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer	120
Qy	376	CAGCCACGGCTGTTCATCATGGTGAAGAAATCAGGCTGCGCTCCCGCTGCGAAACCCCTC	435
Db	121	GlnAlaArgLeuSerSerMetValIylsLysValArgLeuProSerArgCysGluProPro	140
Qy	436	GGAACCACTGTACTGTCTCCGGCTGGGCGACTACCAGAGCCAGATGTCAGCTTTCCTCC	495
Db	141	GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro	160
Qy	496	TCTCACTCATGTGCGTGATGTCAAAGTTCATCTCCCCCAGGACTGCACGAAGTTTAC	555
Db	161	SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTrp	180
Qy	556	AAGGACTTACTGGAAAAATTCATGCTGTGCGCTGGCATCCCGCATCCCAAGAAAAACGCC	615
Db	181	LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla	200
Qy	616	TGCAATGGTGACTCACAGGGGACCGTTGGTGTGCAGAGGTACCTCTCAAGGCTCTGGTGCC	675
Db	201	CysAsnGlyAspSerGlyGlyProLeuValCysargGlyThrLeuGlnGlyLeuValSer	220
Qy	676	TGGGGAATCTTCCCTTGGGCGCAACCCCAATACCCAGGAGTCTACACTCAAGTGTGCAAG	735
Db	221	TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTrpThrGlnValCysLys	240



XX 08-APR-2002; 2002WO-US010824.  
XX 06-APR-2001; 2001US-0281731P.  
PR 06-APR-2001; 2001US-0281732P.  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX Sun Z, Jay G;  
XX WPI; 2003-058520/05.  
DR N-PSDB; ABX10343.  
XX Novel genes which are differentially regulated in prostate cancer, useful  
PT for diagnosing prostate cancer in prostate tissue sample and assessing  
PT therapeutic or preventive intervention in prostate cancer patients.  
XX Claim 1; Page 293-294; 416pp; English.  
XX The invention describes genes (I) which are differentially regulated in  
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
CC sample comprising prostate tissue, which involves determining the number  
CC of target genes which are differentially-regulated in the sample, where  
CC the number is indicative of the probability that the sample comprises  
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
CC intervention in a subject having a prostate cancer, which involves  
CC determining the expression levels in a sample comprising prostate tissue  
CC of target genes which are differentially-regulated in prostate cancer.  
CC Preferably, the expression levels of at least 10 genes are determined.  
CC (I) is also useful for identifying agents that modulate a biological  
CC activity of a polypeptide differentially-regulated in prostate cancer  
CC cells, which involves contacting a polypeptide differentially-regulated  
CC in prostate cancer cells with a test agent under conditions effective for  
CC the test agent to modulate a biological activity of the polypeptide, and  
CC determining whether the test agent modulates the biological activity. (I)  
CC is useful as molecular markers, as drug targets, and for detecting,  
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
CC preventing or treating, determining predisposition to diseases and  
CC conditions especially relating to prostate cancer. (I) and its expression  
CC products are used in the diagnostic test to assay for presence of cancer  
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
CC of cancer, its stage of development, the nature of genetic defect, etc.  
CC The polypeptide encoded by (I) can be used as target for therapy or drug  
CC discovery. (I) can also be used for expressing the polypeptide and thus  
CC for searching specific binding partners of the polypeptide. (I) is useful  
CC in therapeutic applications to treat prostate cancer. The identification  
CC of specific genes, and groups of genes, expressed in pathways  
CC physiologically relevant to prostate cancer permits the definition of  
CC functional and disease pathways and the delineation of targets in these  
CC pathways which are useful in diagnostic, therapeutic, and clinical  
CC applications. This is the amino acid sequence of a protein differentially  
CC regulated in prostate cancer  
XX Sequence 253 AA;  
Alignment Scores:  
Pred. No.: 1.02e-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 6 Gaps: 0  
US-09-905-083A-30 (1-969) x ABU07440 (1-253)  
QY 16 ATGGCAAGATCCCTTCTCTGCGCCCTGCGAGATCTTACTGCTATCTTAGCTTGAAGACT 75  
Db 1 MetAlaArgSerLeuLeuProLeuGlnileLeuLeuLeuLeuLeuLeuLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGCGGTGACAGAGTATTGATGGCGCCCATGTGCAAGAGGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspPyllelleAspGlyAlaProCysAlaArgGlySer 40

QY 136 CACCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTCGGAGGCGTCTCTGTC 195  
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCGCTGGGTGCTCAGTGGCCGCTGCAAGATGAATGATGACACGTCGACCTG 255  
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80  
QY 256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCTCGAAGCTCAATCCGC 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCCATCTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGCCCGCTGCTCATCCATGGTGAAGAAAGTCAAGGCTCGCCCTCCGCTGGAGACCCCT 435  
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGCTACTGCTCCGGCTGGGCACTACACAGAGCCAGATGTGACCTTTCCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTCGTGGTGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAAATTCATGCTGTGGCTGGCATCCCGACTCCAGAAAACGCC 615  
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200  
QY 616 TGCAATGTGACTCAGGGGACCGCTGCTGTGCAGAGTACCCCTGCAAGGTCTGCTGCC 675  
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGGAATTTCCCTTGGCGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC 774  
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253  
RESULT 7  
ABU07471  
ID - ABU07471 standard; protein; 253 AA.  
XX AC ABU07471;  
XX DT 28-JAN-2003 (first entry)  
XX DE Protein differentially regulated in prostate cancer #74.  
XX KW Prostate cancer; gene expression; differential regulation;  
KW molecular marker; drug target; cancer detection; cancer diagnosis;  
KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
XX OS Homo sapiens.  
XX FN WO200281638-A2.  
XX PD 17-OCT-2002.  
XX PF 08-APR-2002; 2002WO-US010824.  
XX PR 06-APR-2001; 2001US-0281731P.  
XX PR 06-APR-2001; 2001US-0281732P.  
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX Sun Z, Jay G;

XX WPI: 2003-058520/05.  
DR N-PSDB; ABX10375.  
XX

Novel genes which are differentially regulated in prostate cancer, useful for diagnosing prostate cancer in prostate tissue sample and assessing therapeutic or preventive intervention in prostate cancer patients.

XX  
PS Claim 1; Page 351; 416pp; English.  
XX

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer

XX  
SQ Sequence 253 AA;

Alignment Scores:  
Pred. No.: 1,028-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 6 Gaps: 0

US-09-905-083A-30 (1-969) x ABU07471 (1-253)

QY 16 ATGGCAAGATCCCTTCCTGCTCCCTCGAGATCTTACTGCTATCTAGCCCTGGAACT 75  
DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20

QY 76 GCAGAGAGAGAGCCAGGGTGACAGATTATTGATGGCGCCCATGTGCAGAGGCTCC 135  
DB 21 AlaGlyGluGluAlaGlnGlyAspIysIleIleAspGlyAlaProCysAlaArgGlySer 40

QY 136 CACCACATGCGAGTGGCTGCTGCTAGTGCAGTACAGTCCACTGGAGGGGCTCTGCTC 195  
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

QY 196 AATGAGCGCTGGTCTCACTGCCGCCCACTGCAGATGAATGATACACCGTGCACCTG 255  
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysIysMetAsnGluTyrThrValHisLeu 80

QY 256 GGCACTGATACGCTGGGCGACAGGAGCTCAGAGGATCAAGGCTCGAAGTCTCATTCGCG 315  
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLeuAlaSerLysSerPheArg 100

QY 316 CACCCCGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGTCAATAGC 375  
DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

QY 376 CAGGCGAGGCTGTTCATCCATGGTGAAGAACTCAGGCTGCCCTCCGCTGCGAACCCTT 435  
DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

QY 436 GGAACCACTCTACTGCTGCTCGGCTGGGGCACTACACAGAGCCAGATGTGACCTTCCC 495  
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

QY 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC 555  
DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

QY 556 AAGGACTTACTGGAAAATTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615  
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

QY 616 TGCATATGCTGACTCAGGGGACCGTGTGTGTGCAGAGTACCTGCAAGGCTGTGTGTTC 675  
DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

QY 676 TGGGGAATCTTCCCTTGGCGCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

QY 736 TTCACCAAGTGGATTAATGACCAATGACCAATGACCAATGACCAATGACCAATGACCA 774  
DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 8  
ABR58471 ID ABR58471 standard; protein; 253 AA.  
XX AC ABR58471;  
XX XX  
DT 07-JUL-2003 (first entry)  
XX XX  
DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.  
XX XX  
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.  
XX XX  
OS Homo sapiens.  
XX XX  
PN WO2003029468-A1.  
XX XX  
PD 10-APR-2003.  
XX XX  
PF 02-OCT-2002; 2002WO-US031467.  
XX XX  
PR 02-OCT-2001; 2001US-0327135P.  
XX XX  
PR 30-MAY-2002; 2002US-0384531P.  
XX XX  
PA (CORI-) CORIXA CORP.  
XX XX  
PI Algate PA, Mannion J;  
XX XX  
DR WPI; 2003-372001/35.  
XX XX  
PT New polynucleotide and polypeptide useful for diagnosing and/or treating  
XX cancer, particularly ovarian cancer, and as a vaccine.  
XX  
PS Claim 2; Page 157-158; 169pp; English.  
XX  
CC The invention relates to a novel isolated polynucleotide. The  
XX polynucleotides of the invention have cytostatic activity, and may have a  
CC use in gene therapy, and in a vaccine. The composition and methods are



CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.  
CC The composition may also be used as a vaccine to prevent cancer. The  
CC present sequence is used in the exemplification of the invention  
XX  
SQ Sequence 253 AA;

Alignment Scores:  
Pred. No.: 1.02e-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 6 Gaps: 0

US-09-905-083A-30 (1-969) x ABR58471 (1-253)

QY 16 ATGGCAAGATCCCTCTCTCTGCGCCCTGCAGATCTTACTGCTATCTTAGCTTGGAACT 75  
DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGGTGCACAGATTATTGATGGCGCCCATGTGCAAGAGCTCC 135  
DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGGCAGGTGGCTGCTGCTAGTGGCAATCAGCTCCAGTGGAGGGCTCTGTGC 195  
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCGTGGGTCTCACTGCGCCGACAGATGAGTGAATGAGTACACCGTGACCTG 255  
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpValHisLeu 80  
QY 256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCCGC 315  
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCGGGTACTCCACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGGCGAGCTGTCTCATGCTGAGAAAGTCAAGCTGCGCTCCGCTGCGAACCCT 435  
DB 121 GlnAlaArgLeuSerMetValLysLysValArgLysProSerArgCysGluProPro 140  
QY 436 GGAACCATCTGTACTGTCTCGGCTGGGGCCTACACAGAGCCAGATGTGACCTTTCCC 495  
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGGGTGGATGTCAGCTCATCTCTCCCGGAGCTGCACGAAGTTTAC 555  
DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAAATTCATGCTGTCGCTGCGCTCCGACTCCCAAGAAAACGCC 615  
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200  
QY 616 TGCATGTGTACTCAGGGGGACCGTTGGTGTGCAGAGGTACCTCGCAAGGTCTGTGTCC 675  
DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACCTTCCCTTGGGGCCCAACCCCAATGACCCAGGAGTCTACATCAAGTGTGCAAG 735  
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGATCCG 774  
DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 9  
ID ADB80484  
XX ADB80484 standard; protein; 253 AA.  
AC ADB80484;

XX 04-DEC-2003 (first entry)  
DT Ovarian cancer-associated protein #24.  
DE  
XX  
KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;  
KW post-operative chemotherapy; radiation therapy; tumour prognosis;  
KW pre-cancerous lesion detection.  
OS Homo sapiens.  
XX WO2002102235-A2.  
FN 27-DEC-2002.  
XX  
PD 18-JUN-2002; 2002WO-US019297.  
XX  
PF 18-JUN-2001; 2001US-0299234P.  
PR 27-AUG-2001; 2001US-0315287P.  
PR 05-SEP-2001; 2001US-0317544P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 12-APR-2002; 2002US-0372246P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA Mack DH, Gish KC;  
XX WPI; 2003-167431/16.  
XX N-PSDB; ADB80483.  
DR  
XX  
PT Detecting an ovarian cancer-associated transcript in a cell from a  
PT patient, comprises contacting a biological sample from the patient with a  
PT polynucleotide that hybridizes to an ovarian cancer gene.  
XX  
PS Claim 13; Page 291; 332pp; English.  
XX  
CC The invention relates to a method of detecting an ovarian cancer-  
CC associated transcript in a cell from a patient, by contacting a  
CC biological sample from the patient with a polynucleotide that selectively  
CC hybridizes to a sequence at least 80% identical to any of one of 80  
CC nucleic acid sequences given in the specification. The method is useful  
CC in diagnosing ovarian cancer and in identifying and using agents and/or  
CC targets that inhibit ovarian cancer. The nucleic acid molecule,  
CC polypeptide and the antibody may also be used in detecting ovarian  
CC cancers, monitoring and early detection of relapse following treatment,  
CC monitoring response to therapy, selecting patients for post-operative  
CC chemotherapy or radiation therapy, in selecting mode of therapy,  
CC determining tumour prognosis, early detection of pre-cancerous lesions,  
CC and as vaccines. This sequence corresponds to one of the proteins used  
CC for the detection method of the invention.  
XX  
SQ Sequence 253 AA;

Alignment Scores:  
Pred. No.: 1.02e-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 7 Gaps: 0  
US-09-905-083A-30 (1-969) x ADB80484 (1-253)  
QY 16 ATGGCAAGATCCCTCTCTGCGCCCTGCAGATCTTACTGCTATCTTAGCTTGGAACT 75  
DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGGTGCACAGATTATTGATGGCGCCCATGTGCAAGAGGTCC 135  
DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGGCAGGTGGCGCTGCTGCTAGTGGCAATCAGCTCCAGTGGAGGGCTCTGTGC 195





QY 496 TCTGACCTCATGTGCTGGATGTCAGCTCAAGCTCACTCCGCCGAGTGCACGAGGTTTAC 555  
Db |||||||  
161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAAATTCATGTGCTGGCGTGGCATCCCGACTCAAGAAAACGCC 615  
Db |||||||  
181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200  
QY 616 TSCAATGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCTGCAAGGTCTGGTGTC 675  
Db |||||||  
201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGGAACCTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db |||||||  
221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGAATAATGACACCATGAAAANGCATCCG 774  
Db |||||||  
241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253  
RESULT 11  
ADN39180  
ID ADN39180 standard; protein; 253 AA.  
XX  
AC ADN39180;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:498.  
XX  
KW Human; differential expression; cancer; angiogenic disorder;  
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
KW inflammatory disease; autoimmune disease;  
KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
KW detection; diagnosis; prognosis; drug screening; drug targeting;  
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
KW vulnery; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO2003042661-A2.  
XX  
PD 22-MAY-2003.  
XX  
PF 13-NOV-2002; 2002WO-US036810.  
XX  
PR 13-NOV-2001; 2001US-0350666P.  
PR 21-NOV-2001; 2001US-0332464P.  
PR 29-NOV-2001; 2001US-0334393P.  
PR 03-DEC-2001; 2001US-0335394P.  
PR 14-DEC-2001; 2001US-0340376P.  
PR 08-JAN-2002; 2002US-0347211P.  
PR 10-JAN-2002; 2002US-0347349P.  
PR 08-FEB-2002; 2002US-0355250P.  
PR 13-FEB-2002; 2002US-0356714P.  
PR 20-FEB-2002; 2002US-0359077P.  
PR 29-MAR-2002; 2002US-0368809P.  
PR 04-APR-2002; 2002US-0370110P.  
PR 12-APR-2002; 2002US-0372246P.  
PR 05-JUN-2002; 2002US-0386614P.  
PR 16-JUL-2002; 2002US-0396839P.  
PR 22-JUL-2002; 2002US-0397775P.  
PR 22-JUL-2002; 2002US-0397845P.  
PR 09-SEP-2002; 2002US-0409450P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
XX WPI: 2003-468649/44.  
DR N-PSDB; ADN39179.

XX Determining the presence or absence of a pathological cell in a patient,  
PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
PT a nucleic acid in a biological sample.  
XX  
FS Claim 12; SEQ ID NO 498; 1385pp; English.  
XX  
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
CC whose expression is upregulated or downregulated in specific cancers or  
CC other diseases such as angiogenic or fibrotic disorders, and to methods  
CC of determining the presence or absence of a pathological cell in a  
CC patient by detecting a nucleic acid at least 80% identical to those of  
CC the invention or by detecting a polypeptide of the invention. The  
CC invention also relates to expression vectors and host cells comprising a  
CC nucleic acid of the invention; antibodies which specifically bind a  
CC polypeptide of the invention; use of such antibodies for drug targeting;  
CC and methods of screening for modulators of activity or expression of the  
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
CC antibodies and methods are useful for diagnosing, prognosing and treating  
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
CC neovascularisation syndromes, scarring and uterine fibroids. They may  
CC also be useful in wound healing and in contraception. The present  
CC sequence represents a polypeptide of the invention.  
XX  
SQ Sequence 253 AA;  
Alignment Scores:  
Pred. No.: 1.02e-134 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 7 Gaps: 0  
US-09-905-083A-30 (1-969) x ADN39180 (1-253)  
QY 16 ATGGCAAGATCCCTTCTCTGCCCCCTGCAGATCTTACTGTCTATCTTAGCCTTGGAACT 75  
Db |||||||  
1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAAGAGCCAGGGTGACAGATTATTGATGGCGCCCATGTGCAAGAGGTCC 135  
Db |||||||  
21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCCATGGCAGGTGGCCCTGCTCAGTGGAATCAGCTCCACCTCGGAGGGCTCTGTC 195  
Db |||||||  
41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCGCTGGTGTCTACTGCGCCCTGCAAGATGAATGATGATACCGTGCACCTG 255  
Db |||||||  
61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80  
QY 256 GGCAGTGATACGCTGGCGGACAGGAGCTCAGAGGATCAAGGCTCGAAGTCAATCCGC 315  
Db |||||||  
81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
Db |||||||  
101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGGCCAGGCTGTCTATCCATGGTGAAGAAAGTACAGGCTGCCCTCCGCTCGGAACCCCT 435  
Db |||||||  
121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTGGGCTGGGGCACTACACAGAGCCAGATGTGACCTTTCCC 495  
Db |||||||  
141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGAGCTCAGCAAGGTTTAC 555  
Db |||||||  
161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

QY 556 AAGGACTTACTGAAATTCATGCTGTGGCTGGCATCCCGACTCCCAAGAAACGCC 615  
 DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
 QY 616 TGCATGTGACTCAGGGGACCGTGTGTGTCAGAGGTACCTGCAAGTCTGTGTCC 675  
 DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
 QY 676 TGGGAACTTCCCTTGGCGCCCAACCCATGACCCAGGAGTCTACACTCAAGTGTCAAG 735  
 DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValThrGlnValCysLys 240  
 QY 736 TTCACCAAGTGTGATAAATGACACCACTGAAAGACATCGC 774  
 DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

## RESULT 12

ADL06515

ID ADL06515 standard; protein; 253 AA.

XX AC ADL06515;

XX DT 20-MAY-2004 (first entry)

XX DE Human tumour-associated antigenic target (TAT) polypeptide #14.

XX KW Human; tumour-associated antigenic target; TAT; cell death; tumour;

XX KW cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2004016225-A2.

XX PD 26-FEB-2004.

XX PF 19-AUG-2003; 2003WO-US025892.

XX PR 19-AUG-2002; 2002US-0404809P.

XX PR 21-AUG-2002; 2002US-0405645P.

XX PR 23-SEP-2002; 2002US-0413192P.

XX PR 15-OCT-2002; 2002US-0419008P.

XX PR 15-NOV-2002; 2002US-0426847P.

XX PR 02-JUL-2003; 2003US-0484959P.

XX PA (GETH) GENENTECH INC.

XX PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

XX PI Spencer SD, Wu TD, Zhang Z;

XX XX WPI: 2004-257144/24.

XX DR N-PSDB; ADL06435.

XX PT New antibody that binds to a tumor-associated antigenic target (TAT)

XX PT polypeptide, useful for preparing a composition for diagnosing or

XX PT treating cancer.

XX PS Claim 2; SEQ ID NO 95; 319pp; English.

XX CC The present invention relates to the isolation of human tumour-associated  
 CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also  
 CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is  
 CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a  
 CC humanised antibody. It is conjugated to a growth inhibitory agent. It is  
 CC produced in bacteria or in CHO cells and induces death of a cell to which  
 CC it binds. The antibody is useful for preparing a composition for  
 CC diagnosing or treating tumours and cancer. The present sequence  
 CC represents a human TAT polypeptide of the invention.

XX SQ Sequence 253 AA;

Alignment Scores:

Pred. No.: 1,02e-134 Length: 253

Score: 1364.00 Matches: 253

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 76.63% Indels: 0  
 DB: 8 Gaps: 0

US-09-905-083A-30 (1-969) x ADL06515 (1-253)

QY 16 ATGGCAGATCCCTCTCTCTGCTGGCTGGCATCCCGACTCTTACTGTATCTTACCTTGGAACT 75  
 DB 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20  
 QY 76 GCAGAGAGAAGAGCCAGGGTGCACAAAGATTATTGATGGCGCCCATGTCGCAAGAGCTCC 135  
 DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
 QY 136 CACCATGGCAGGTGGCCCTCTCTCAGTGGCAATCAGCTCCACTGCGAGGGGTCTCTGTCT 195  
 DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
 QY 196 AATGAGCGCTGGGTGCTCAGTCCGCCACACGCAAGATGAATGATGACACCGTGCACCTG 255  
 DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluIuThrValHisLeu 80  
 QY 256 GGCAGTATACGCTGGCGCACAGGAGCTCAGAGGATCAAGGCTCGAAGCTCATTTCCGC 315  
 DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
 QY 316 CACCCCGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
 DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
 QY 376 CAGCCCGCTGCTCATCCATGGTGAAGAAAGTCAAGGCTGCGCTCCCGCTGCGAACCCCT 435  
 DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
 QY 436 GGAACCACTCTACTGTCTCCGGCTGGGGGACTACACAGACCCAGATGTGCACCTTTCC 495  
 DB 141 GlyThrThrCysThrValSerGlyTyrGlyThrThrSerProAspValThrPhePro 160  
 QY 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC 555  
 DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
 QY 556 AAGCACTTACTGGAAAATTCATGCTGCGCTGGCATCCCGACTCCCAAGAAAACGCC 615  
 DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
 QY 616 TGCATGTGACTCAGGGGACCGTGTGTGTGTCAGAGGTACCTGCAAGGTCTGTGTCTCC 675  
 DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
 QY 676 TGGGAACTTTTCCCTTGGCGCCAAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
 DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
 QY 736 TTCACCAAGTGTGATAAATGACACCACTGAAAGACATCGC 774  
 DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

## RESULT 13

ADN04182

ID ADN04182 standard; protein; 253 AA.

XX AC ADN04182;

XX DT 01-JUL-2004 (first entry)

XX DE Antipsoriatic protein sequence #286.

XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.

XX OS Homo sapiens.



## Alignment Scores:

Pred. No.: 1,02e-134 Length: 253  
 Score: 1364.00 Matches: 253  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 76.63% Indels: 0  
 DB: 8 Gaps: 0

US-09-905-083A-30 (1-969) x ADR72880 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTGCTGCCCTCGAGATCTTACTGCTATCTTACTGCTTGTGAAACT 75  
 DB 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuLeuSerLeuAlaLeuGluThr 20  
 QY 76 GCAGGAGAGAGCCAGGTCACAGATTATTGATGGCGCCCATGTCAGAGGCTCC 135  
 DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
 QY 136 CACCATGCGAGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGGGAGGCGCTCTGTGTC 195  
 DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
 QY 196 ANTGAGCGTGGTGTCTACTGCCGCCCACTGCAAGATGAATGATACCGTGCACCTG 255  
 DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80  
 QY 256 GCGAGTATACCTGGCGCAGGAGCTCAGAGATCAAGGCTCGAAGTCATTCGCC 315  
 DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
 QY 316 CACCCCGGCTACTCCACAGACCATGTTAATGACCTCATCTCGTGAAGTCAATAGC 375  
 DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
 QY 376 CAGGCCAGCTCTCATCCATGTGTAAGAAAGTCAGGCTGCCCTCCGCTCGCAACCCCT 435  
 DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
 QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGGCCAGATGTGACCTTCCC 495  
 DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
 QY 496 TCTGACCTCATGTGCTGATGTCAAGCTCATCTCCCCCAGGACTCACAGAGGTTTAC 555  
 DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValThr 180  
 QY 556 AAGGACTTACTGGAATAATTCATGTGCTGCTGCCATCCCGACTCCCAAGAAAACGCC 615  
 DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200  
 QY 616 TCCATGTGACTCAGGGGACCGTGTGTGTCAGAGGTACCTCGAAGGTCTGTGTGCC 675  
 DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
 QY 676 TGGGGAACCTTCCCTTTGGGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
 DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValThrGlnValCysLys 240  
 QY 736 TTCACCAAGTGTGATAAATGACACCATGATAAAGACATCGC 774  
 DB 241 PheThrLysTrpIleAspAspThrMetLysLysHisArg 253

## RESULT 15

ADA05732

ID ADA05732 standard; protein; 250 AA.

XX AC ADA05732;

XX 06-NOV-2003 (first entry)

XX Human NOV18a protein SEQ ID NO:92.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilepaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 XX WO2003029424-A2.  
 XX 10-APR-2003.  
 XX 02-OCT-2002; 2002WO-US031373.  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339286P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.  
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Raschell L, Stone DU, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg MB, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI: 2003-381626/36.  
 XX N-FSDB; ADA05731.  
 DR New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 DR preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 PT  
 XX Claim 1; Page 169-170; 586pp; English.  
 PS  
 XX The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX  
SQ

Sequence 250 AA;

Alignment Scores:  
Pred. NO.: 3.07e-133 Length: 250  
Score: 1350.00 Matches: 250  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 75.84% Indels: 0  
DB: 6 Gaps: 0

US-09-905-083A-30 (1-969) x ADA05732 (1-250)

QY	25	TCCTCTCTCCCTCCCTCAGATCTTACTGTATCTTAGCTTGGAACTGCAGAGAA	84
DB	1	SerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThrAlaGlyGlu	20
QY	85	GAAGCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCCAACCATGG	144
DB	21	GluAlaGlnGlyAspLeuAlaIleAspGlyAlaProCysAlaArgGlySerHisProTrp	40
QY	145	CAGGTGGCCCTGCTCAGTGGGCAATCAGCTCCACTGCGGAGGCGTCTGGTCAATGAGCGC	204
DB	41	GlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArg	60
QY	205	TGGGTGCTCACTGCGCCGCACTGCAGATGAATGAGTACACCGTGCACTGGGCAATGAT	264
DB	61	TrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAsp	80
QY	265	ACGTGGGCGCAGAGAGCTCAGAGGATCAAGGCTCGAAGTCAATTCGCGCACCCCGGC	324
DB	81	ThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGly	100
QY	325	TACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGCCAGG	384
DB	101	TyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArg	120
QY	385	CTGTCTCATGTGGTGAAGAGTCAAGGCTGCGCTCCCTCGGAAACCCCTGGAAACACC	444
DB	121	LeuSerSerMetValLysValArgLeuProSerArgCysGluProProGlyThrThr	140
QY	445	TGTACTGTCTCGGTGGGCGACTACACAGCCAGATGTACCTTTCCCTCTGACCTC	504
DB	141	CysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeu	160
QY	505	ATGTGCGTGGATGTCAAGCTCATCTCCCGCCAGGACTGCAGAGGTTTACAGGACTTA	564
DB	161	MetCysValAspValLysLeuIleSerProGlnAspCysThrLysValLysAspLeu	180

QY	565	CTGGAATAATTCATGCTGTGGCTGGCATCCCCGACTCCCAAGAAAAACGCTGCAATGGT	624
DB	181	LeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGly	200
QY	625	GACTCAGGGGAGACCGTTGGTGTGCAGAGGTACCTCGCAAGGTCTGGTGTCTCTGGGGA	684
DB	201	AspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThr	220
QY	685	TTCCCTTGGCGCCCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTCAAGTTCACCA	744
DB	221	PheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLys	240
QY	745	TGGATAAATGACACCATCAAAAAAGCATCGC	774
DB	241	TrpIleAsnAspThrMetLysLysHisArg	250

Search completed: July 14, 2005, 17:25:05  
Job time : 177.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 14, 2005, 17:16:29 ; Search time 37.5 seconds

(without alignments)

3857.864 Million cell updates/sec

Title: US-09-905-083A-30

Perfect score: 1780

Sequence: 1 ggatttcgggctccatggc.....aagaacacacaaacctcag 969

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+npz.model -DEV=xlp  
-Q=/cgm2\_1/USPTO\_spool\_p/US09905083/runat\_14072005\_140354\_10640/app\_query.fasta\_1.1159  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=opt -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09905083@cgm2\_1\_46 @runat\_14072005\_140354\_10640 -NCFU=6 -ICPU=3  
-NO WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:

- 1: /cgm2\_6/ptodata/1/iaa/5A.COMB.pcp:\*\*
- 2: /cgm2\_6/ptodata/1/iaa/5B.COMB.pcp:\*\*
- 3: /cgm2\_6/ptodata/1/iaa/6A.COMB.pcp:\*\*
- 4: /cgm2\_6/ptodata/1/iaa/6B.COMB.pcp:\*\*
- 5: /cgm2\_6/ptodata/1/iaa/6CTUS.COMB.pcp:\*\*
- 6: /cgm2\_6/ptodata/1/iaa/backfiles.pcp:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1371	77.0	265	4	US-09-949-016-7716
2	1364	76.6	253	2	US-08-557-146-2
3	1364	76.6	253	2	US-08-824-874-3
4	1364	76.6	253	2	US-09-154-344-2
5	1364	76.6	253	3	US-08-930-188-2
6	1364	76.6	253	3	US-09-210-084-3
7	1364	76.6	253	4	US-09-764-762-3
8	1364	76.6	253	5	PCT-US96-04294-2
9	1235	69.4	225	2	US-09-027-337-4
10	1235	69.4	225	4	US-09-644-600-4
11	1235	69.4	225	4	US-09-654-600A-4
12	1230	69.1	224	3	US-08-944-483-33

13	1215.5	68.3	225	2	US-08-557-146-12	Sequence 12, Appl
14	1215.5	68.3	225	2	US-09-154-344-12	Sequence 12, Appl
15	829	46.6	154	3	US-09-261-416-7	Sequence 7, Appl
16	773	43.4	144	4	US-08-618-259-4	Sequence 4, Appl
17	614	34.5	260	3	US-09-025-059-3	Sequence 8, Appl
18	614	34.5	260	4	US-09-618-259-8	Sequence 8, Appl
19	612.5	34.4	270	4	US-09-949-016-7712	Sequence 7712, Ap
20	612.5	34.4	293	4	US-09-509-908-2	Sequence 2, Appl
21	582.5	32.7	260	3	US-09-070-526-2	Sequence 2, Appl
22	582.5	32.7	260	4	US-08-618-259-7	Sequence 7, Appl
23	582.5	32.7	290	4	US-09-949-016-8166	Sequence 8166, Ap
24	580.5	32.6	260	3	US-09-008-271A-7	Sequence 7, Appl
25	573.5	32.2	282	3	US-09-025-059-1	Sequence 1, Appl
26	571.5	32.1	250	4	US-09-205-258-427	Sequence 427, App
27	568.5	31.9	288	4	US-09-386-642-13	Sequence 13, Appl
28	567.5	31.9	254	3	US-09-439-313-525	Sequence 525, App
29	567.5	31.9	254	4	US-09-636-215-525	Sequence 525, App
30	567.5	31.9	254	4	US-09-685-166A-525	Sequence 525, App
31	567.5	31.9	254	4	US-09-679-426-525	Sequence 525, App
32	567.5	31.9	254	4	US-09-759-143-525	Sequence 525, App
33	567.5	31.9	254	4	US-09-651-236-525	Sequence 525, App
34	561	31.5	246	2	US-08-978-404B-44	Sequence 44, Appl
35	559.5	31.4	449	4	US-09-636-215-617	Sequence 617, App
36	559.5	31.4	449	4	US-09-685-166A-617	Sequence 617, App
37	559.5	31.4	449	4	US-09-679-426-617	Sequence 617, App
38	559.5	31.4	449	4	US-09-759-143-617	Sequence 617, App
39	559.5	31.4	449	4	US-09-651-236-617	Sequence 617, App
40	558	31.3	248	1	US-08-744-026-1	Sequence 1, Appl
41	558	31.3	248	2	US-09-102-732-1	Sequence 1, Appl
42	558	31.3	248	3	US-09-261-767-1	Sequence 1, Appl
43	558	31.3	268	2	US-08-824-874-1	Sequence 1, Appl
44	558	31.3	268	3	US-09-210-084-1	Sequence 1, Appl
45	558	31.3	268	4	US-09-764-762-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-7716  
; Sequence 7716, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C0001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7716  
; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7716

Alignment Scores:  
Pred. No.: 9.84e-131 Length: 265  
Score: 1371.00 Matches: 254  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.61% Mismatches: 0  
Query Match: 77.02% Indels: 0  
DB: 4 Gaps: 0

US-09-905-083A-30 (1-969) x US-09-949-016-7716 (1-265)

Oy 10 GGCTCCATGGCAAGATCCCTCTCTCCCTCCGAGATCTTACTGCTATCTTAGCCTTG 69



```
Db 11 GlyThrMetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeu 30
Qy 70 GAAACTGCAGGAGAGAGCCAGAGGTGACAGATTATTGATGGCGCCCATGTGCAAGA 129
Db 31 GluThrAlaGlyGluGlnAlaGlnGlyAspIleAspGlyAlaProCysAlaArg 50
Qy 130 GCTCTCCACCCATGCGAGTGGCCCTGCTCAGTGGCAATCAAGTCCACTCGGAGGCGTC 189
Db 51 GlySerHisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyVal 70
Qy 190 CTGGTCAATGAGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGATACCGGTG 249
Db 71 LeuValAsnGlnArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrVal 90
Qy 250 CACTTGGGCGAGTATACGTGGCGACAGGAGGCTCAGAGGATCAAGGCTCGAAGTCA 309
Db 91 HisLeuGlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSer 110
Qy 310 TTCCGCCACCCGGCTACTCCACAGACCCCATGTTAATGACCTCATCTCGTGAAGCTC 369
Db 111 PheArgHisProGlyTy-SerThrGlnThrHisValAsnAspLeuMetLeuValLysLeu 130
Qy 370 AATACCCAGGCGAGCTGCTATCCATGTTGAAGAAAGTCAAGCTCCCTCCGCTGCGAA 429
Db 131 AsnSerGlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGlu 150
Qy 430 CCCCTTGAACACCACTGTACTGCTCCGCTGGGGCACTACCACGAGCCAGATGTGACC 489
Db 151 ProProGlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThr 170
Qy 490 TTTCCCTCTGACCTCATGTGCGTGGATGTCAGAGTTCATCTCCCCCAGGACTGCACAG 549
Db 171 PheProSerAspLeuMetCysValAspValLysLeuLysSerProGlnAspCysThrLys 190
Qy 550 GTTTACAGGACTTACTGAAATTCATGCTGCTGGCTGGCATCCCCGACTCCAGAAA 609
Db 191 ValTyLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLys 210
Qy 610 AACGCTGCAATGTTGACTTCAGGGGGACCGTTGGTGTGCAGAGGTACCTGCAAGTCTG 669
Db 211 AsnAlaCysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeu 230
Qy 670 GTGTCTCGGGAACTTCCCTTTCGCGCCCAACCAATGACCCAGGAGTCTACACTCAAGTG 729
Db 231 ValSerTrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyThrGlnVal 250
Qy 730 TGCAGTTTCAACCAAGTGTATTAATGACACCATGAAAAGCATCGC 774
Db 251 CysLysPheThrLysTrpIleAsnAspThrMetLysLysHisArg 265
```

## RESULT 2

```
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-557-146-2
```

## Alignment Scores:

```
Pred. No.: 4.97e-130 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
DB: 2 Gaps: 0
```

US-09-905-083A-30 (1-969) x US-08-557-146-2 (1-253)

```
Qy 16 ATGCACAGATCCCTTCCTGCGCCCTGAGATCTTACTGTCTATCTTACTGCTTGGAAACT 75
Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20
Qy 76 GCAGGAGAAGAGCCAGGGTGACAAGATTATTCAGCGCCCATGTCGCAAGAGGCTCC 135
Db 21 AlaGlyGluGlnAlaGlnGlyAspLysIleLeuAspGlyAlaProCysAlaArgGlySer 40
Qy 136 CACCCTATGCGAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTCGGAGGCGTCTGTGTC 195
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60
Qy 196 AATGAGCGCTGGGTGCTCACTGCGCCCTGCAAGATGAATGATGACCGTGCACCTG 255
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyThrValHisLeu 80
Qy 256 GGCAGTGTATAGCTGGCGCACAGGAGACTCAGAGGATCAAGGCTCGAAGTCAATTCGC 315
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
Qy 316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375
Db 101 HisProGlyTySerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
Qy 376 CAGCCAGGCTGTCTATCCATGCTGGAAGAAAGTCAAGGCTGCGCTCCCTCCGCTGGAACCCCT 435
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
Qy 436 GGAACCACTGTACTGCTCCGCGTGGCGCACTACACAGAGCCAGAGTGTGACCTTTTCCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
Qy 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGAGTGTGACGAAAGTTTAC 555
Db 161 SerAspLeuMetCysValAspValLysLeuLysSerProGlnAspCysThrLysValTy 180
Qy 556 AAGCACTTACTGGAATAATTCATGCTGTGGCTGGCATCCCCGACTCAAGAAAACGCC 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
Qy 616 TGCATGTGCTACTCAGGGGACCGTGTGTGTGACAGAGTACCTTCGCAAGGCTCTGTGTCC 675
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
```

QY 676 TGGGGAACCTTCCCTGCGGCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAspProGlyValThrGlnValCyslys 240  
QY 736 TTCACCAAGTGAATAAATGACACCAATGAAAGAGTCCG 774  
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 3  
US-08-824-874-3  
; Sequence 3, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: S14  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 532504  
US-08-824-874-3

Alignment Scores:  
Pred. No.: 4,978-130 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 2 Gaps: 0

US-09-905-083A-30 (1-969) x US-08-824-874-3 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTGCGCCCTGAGATCTTACTGCTATCTAGCCCTGGAAGT 75  
Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGTCACAGATTATTGATGGCGCCCATGTCAGAGGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCTGCTC 195

Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCCCTGGGTGCTCACTGCCGCCCTGCAAGATGAATGATGATGACCGTGCACCTG 255  
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80  
QY 256 GGCAGTGATACGCTGGCGGACAGGAGCTCAGAGGATCAAGGCCCTCAAGTCATTCCGC 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTTACTCCACACAGACCCATCTTAATGACCTCATGCTCGTGAAGTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGCCAGGCTGTCATCCATGCTGAAGAAAGTCAAGCTGCCCTCCCGCTGCCAAGCCCTC 435  
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCGGCTGGGGGCACTACACAGGCCCAAGATGTGACCTTCCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGCTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGTGTTCAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAAATTCATGCTGTGCTGCGCTGCCATCCCGACTCAAGAAAAACGCC 615  
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
QY 616 TGCAATGCTGACTCAGGGGGACCGTGTGTGTCAGAGGTACCTGCAAGTGTGTGTC 675  
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACTTCCCTTGGGGCAACCCCAATGACCCAGGAGTACCTGCAAGTGTGTGTC 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGCATCCG 774  
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 4  
US-09-134-344-2  
; Sequence 2, Application US/09154344  
; Patent No. 5981256  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/154,344  
; FILING DATE: 16-SEP-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION:



```

; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-09-210-084-3

Alignment Scores:
Pred. No.: 4,97e-130 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
Gaps: 3

US-09-905-083A-30 (1-969) x US-09-210-084-3 (1-253)
QY 16 ATGGCAAGATCCCTTCTCTGCCCCGCTGCAGATCTTACTGTCTATCTTAGCCTTGGAACT 75
Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20
QY 76 GCAGAGAAAGAGCCAGGGTGACAGATTATTGATGGGCCCCCATGTGCAAGAGCTCC 135
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATCGCAGGTGGCCCTCTCAGTGGCAATCAGCTCCACCTGGGAGGCGTCTGGTC 195
Db 41 HisProThrGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCCTGGGTGCTCCTCCTGCCCCCAGTGCAGATGAATGAGTACACCTGCACCTG 255
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80
QY 256 GCGAGTGATACGCTGGGGCAGAGGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACACAGACCCATGTTAATGACCTCATCTGCTGTAAGCTCAATAGC 375
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCCAGGCTGCTATCCATGCTGAAGAAAGTCAAGGCTGCCCTCCCGTGCAGAACCCCT 435
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCAGATGTGACCTTTCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATCTGCTGGTGGATGCTCAAGCTCATCTCCCCCAGGACTGCACGAAGTTTAC 555
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValThr 180
QY 556 AAGGACTTACTGGAATAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TGCATGTGTGACTCAGGGGGACCGTTGTTGTGTGAGAGGTACCTGCAAGGTCTGTGTGCC 675
Db 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGGAACCTTCCCTTGGGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValThrGlnValCysLys 240
QY 736 TTCACCAAGTGTGATAAATGACACCATGAAAGCAATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

```

## RESULT 6

```

US-09-210-084-3
; Sequence 3, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```

```

; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-09-210-084-3

Alignment Scores:
Pred. No.: 4,97e-130 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
Gaps: 3

US-09-905-083A-30 (1-969) x US-09-210-084-3 (1-253)
QY 16 ATGGCAAGATCCCTTCTCTGCCCCGCTGCAGATCTTACTGTCTATCTTAGCCTTGGAACT 75
Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20
QY 76 GCAGAGAAAGAGCCAGGGTGACAGATTATTGATGGGCCCCCATGTGCAAGAGCTCC 135
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATCGCAGGTGGCCCTCTCAGTGGCAATCAGCTCCACCTGGGAGGCGTCTGGTC 195
Db 41 HisProThrGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCCTGGGTGCTCCTCCTGCCCCCAGTGCAGATGAATGAGTACACCTGCACCTG 255
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80
QY 256 GCGAGTGATACGCTGGGGCAGAGGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACACAGACCCATGTTAATGACCTCATCTGCTGTAAGCTCAATAGC 375
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCCAGGCTGCTATCCATGCTGAAGAAAGTCAAGGCTGCCCTCCCGTGCAGAACCCCT 435
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCAGATGTGACCTTTCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATCTGCTGGTGGATGCTCAAGCTCATCTCCCCCAGGACTGCACGAAGTTTAC 555
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValThr 180
QY 556 AAGGACTTACTGGAATAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TGCATGTGTGACTCAGGGGGACCGTTGTTGTGTGAGAGGTACCTGCAAGGTCTGTGTGCC 675
Db 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGGAACCTTCCCTTGGGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValThrGlnValCysLys 240
QY 736 TTCACCAAGTGTGATAAATGACACCATGAAAGCAATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

```



! INFORMATION FOR SEQ ID NO: 2:  
! SEQUENCE CHARACTERISTICS:  
! LENGTH: 253 amino acids  
! TYPE: amino acid  
! STRANDEDNESS: single  
! TOPOLOGY: linear  
! MOLECULE TYPE: protein  
! PCT-US96-04294-2

Alignment Scores:  
Pred. No.: 4,97e-130 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: Gaps: 0

US-09-905-083A-30 (1-969) x PCT-US96-04294-2 (1-253)

```
QY 16 ATGGCAAGATCCCTTCTCTGCGCCCTGCGATCTTACTGCTATCTTAGCCTTGGAACT 75
DB 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20
QY 76 GCAGGAGAGAGCCAGGCTGACAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC 135
DB 21 AlaGlyGluGlnGlnGlyAspIleIleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCAATGCGAGTGCCCTGCTCACTGGCAATCACTCCACTGCGGAGCGCTCTGTGTC 195
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60
QY 196 ATGAGCGCTGGTGCTCACTGCGCCCTCACTGCAAGATGAATGAGTACACCGTGCACTG 255
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80
QY 256 GCGAGTGATACCGTGGCGACAGGAGCTCAGAGATCAAGGCTCGAAGTCATTCCGC 315
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCCGGCTACTCCACAGACCCCATGTTAATGACCTCATCTCGTGAAGCTCAATAGC 375
DB 101 HisProGlyTy-SerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCGAGGCTGCTCACTGCTGAGAAAGTCAGGCTGCGCTCCCTCGCGTCCGAAACCCCT 435
DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGCTGGGGCACTACCACGAGCCAGATGTGACCTTTCC 495
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATGTGCGTGATGTCAGGCTCATCTCCCCCAGGACTGCACGAAAGTTTAC 555
DB 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTy 180
QY 556 AAGGACTTACTGAAATTCATGCTGCGCTGCGATCCCGACTCCGAAAGAAACGCC 615
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TCAATGGTGACTCAGGGGACCGTGGGTGTCAGAGGTACCTCGAAGGTCTGTGTGCC 675
DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGGAACCTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyThrGlnValCysLys 240
QY 736 TTCACCAAGTGAATAATGACACCATGAAAGATCGC 774
DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 9

US-09-027-337-4

! Sequence 4, Application US/09027337B  
! Patent No. 5972616  
! GENERAL INFORMATION:  
! APPLICANT: O'Brien, Timothy J.  
! APPLICANT: Tanimoto, Hiroto  
! TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
! TITLE OF INVENTION: Breast and Ovarian Carcinomas  
! FILE REFERENCE: D6064  
! CURRENT APPLICATION NUMBER: US/09/027,337B  
! CURRENT FILING DATE: 1998-02-20  
! NUMBER OF SEQ ID NOS: 13  
! SEQ ID NO 4  
! LENGTH: 225  
! TYPE: PRT  
! ORGANISM: Unknown  
! FEATURE:  
! OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to  
! OTHER INFORMATION: similar domain in TADG-15

US-09-027-337-4

Alignment Scores:  
Pred. No.: 6.11e-117 Length: 225  
Score: 1235.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 69.38% Indels: 0  
DB: Gaps: 0

US-09-905-083A-30 (1-969) x US-09-027-337-4 (1-225)

```
QY 100 AAGATTATTGATGGCGCCCATGTGCAAGAGCTCCACCCATGCGAGTGGCCCTGTC 159
DB 1 LysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeu 20
QY 160 AGTGCAATCACTCCACTGCGGAGCGCTCTGTTCAATGAGCGCTGGGTGCTCACTGCC 219
DB 21 SerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40
QY 220 GCGCACTCAAGATGAATGAGTACACCGTGCACCTGGGAGTGATACGTTGGGCGACAGG 279
DB 41 AlaHisCysLysMetAsnGluTrpThrValHisLeuGlySerAspThrLeuGlyAspArg 60
QY 280 AGAGCTCAGAGATCAAGGCTCGAAGTCATTCCGCCACCCCGCTACTCCACACAGACC 339
DB 61 ArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTy-SerThrGlnThr 80
QY 340 CATGTTAATGACCTCATCTGCTGAAGCTCAATAGCCAGGCGAGCTGTCAATCCATGGTG 399
DB 81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100
QY 400 AAGAAAGTCAGGCTGCCCTCCGCTGCGAACCCCTCGAACCACTGTACTGTCTCCGCGC 459
DB 101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120
QY 460 TGGGCACTACACGAGCCAGATGTGACCTTCCCTCTGACCTCATGTGCGTGGATGTC 519
DB 121 TrpGlyThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal 140
QY 520 AAGCTCATCTCCCCCAGGACTGCACGAAAGGTTTCAAGAGGACTTACTGAAAAATTCATG 579
DB 141 LysLeuIleSerProGlnAspCysThrLysValTyLysAspLeuLeuGluAsnSerMet 160
QY 580 CTGTGCGTGCATCCCGGACTCCCAAGAAAAACCGCTGCAATGGTGACTCAGGGGACCG 639
DB 161 LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
QY 640 TTGGTGTGACAGGTTACCTTCAAGGTTGGTGTCTCTGGGNACTTTCCCTTCCGCGCNA 699
DB 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln 200
QY 700 CCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCCCAAGTGGATAAATCACACC 759
DB 201 ProAsnAspProGlyValTyThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220
```





```

QY 460 TGGGGCACTTACCACGAGCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGTGATGC 519
DB 121 TTPGlyThrThrThrSerProaspValThrPheProSerAspLeuMetCysValaspVal 140
QY 520 AAGCTCATCTCCCCCAGGACTGCACGAGGTTTACAAGGACTTACTGGAATAATTCATG 579
DB 141 LysLeuIleSerProGlnAspCysThrLysValTyrLysAspLeuGluAsnSerMet 160
QY 580 CTGTGCGCTGGCATCCCGCTCCAGAAAGGCTGCAATGGTGACTCAGGGGACCG 639
DB 161 LeuCysAlaGlyLeuProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
QY 640 TTGGTGTGCAGAGGTACCTCGCAAGCTCTGGTGTCTGGGGAACCTTTCCTTGGCGCCAA 699
DB 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln 200
QY 700 CCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTTCCCAAGTGGATAAATGACACC 759
DB 201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220
QY 760 ATGAAAAGCATCGC 774
DB 221 MetLysLysHisArg 225

```

## RESULT 12

```

US-08-944-483-33
; Sequence 33, Application US/08944483
; Patent No. 6232456

```

```

; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE

```

```

; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.

```

```

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

```

```

; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

```

```

; NUMBER OF SEQUENCES: 76

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL

```

```

; COUNTRY: USA
; ZIP: 60064-3500

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette

```

```

; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: DOS

```

```

; SOFTWARE: FastSeq for Windows Version 2.0

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/944,483

```

```

; FILING DATE:

```

```

; CLASSIFICATION: 424

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:

```

```

; FILING DATE:

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Becker, Cheryl L.

```

```

; REGISTRATION NUMBER: 35,441

```

```

; REFERENCE/DOCKET NUMBER: 6183.US.01

```

```

; TELEPHONE: 847/935-1729

```

```

; TELEFAX: 847/938-2623

```

```

; TELEX:

```

```

; INFORMATION FOR SEQ ID NO: 33:

```

```

; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 224 amino acids

```

```

; TYPE: amino acid

```

```

;

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-33

```

## Alignment Scores:

```

Pred. No.: 1,96e-116 Length: 224
Score: 1230.00 Matches: 224
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.10% Indels: 0
DB: 3 Gaps: 0

```

```

US-09-905-083A-30 (1-969) x US-08-944-483-33 (1-224)

```

```

QY 103 ATTATTGATGGCGCCCATGTGCAAGAGCTCCACCCCATGGCAGGTGGCCCTGCTCAGT 162
DB 1 IleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeuSer 20
QY 163 GGCATCAGCTCCACTCGCGAGGCGTCTGCTCAATGAGCGCTGGTGTCTCACTGCCGCC 222
DB 21 GlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAlaAla 40
QY 223 CACTGCAAGATGAATGAGTACACCGCTGCACCTGGCGAGTATAGCTGGCGCACAGGAGA 282
DB 41 HisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyAspArg 60
QY 283 GCTCAGAGGATCAAGGCTCGAAGTCAATTCGCGCACCCCGGCTACTCCACACAGACCCAT 342
DB 61 AlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyrSerThrGlnThrHis 80
QY 343 GTTAATGACCTCATGCTCGTGAAGTCAATAGCCAGGCGAGCTGTCAATCCATGCTGAAG 402
DB 81 ValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerMetValLys 100
QY 403 AAGTCAAGCTGCGCTCCCGCTGCGAAACCCCTGGAACACCTGTACTGTCTCCGGCTGG 462
DB 101 LysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGlyTrp 120
QY 463 GGCATCACCAGCGCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAG 522
DB 121 GlyThrThrThrSerProAspValThrPheProSerAspLeuMetCysValAspValLys 140
QY 523 CTGATCTCCCCCAGGACTGCACGAGGTTTACAGGACTTACTGGAAAATTCATGCTG 582
DB 141 LeuIleSerProGlnAspCysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeu 160
QY 583 TGCCTGCGATCCCGGACTCCCAAGAAAACCGCTGCAATGTGTGACTCAGGGGGACCGTGT 642
DB 161 CysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyProLeu 180
QY 643 GTGTGCAGAGGTACCTTCGCAAGGTCTGGTGTCTCTGGGGAACCTTTCCCTTGGCGCAACCC 702
DB 181 ValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGlnPro 200
QY 703 AATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCCACAGTGGATAATGACACCATG 762
DB 201 AsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThrMet 220
QY 763 AAAAAGCATCGC 774
DB 221 LysLysHisArg 224

```

## RESULT 13

```

US-08-557-146-12

```

```

; Sequence 12, Application US/08557146

```

```

; Patent No. 5834290

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Egelrud, Torbjorn

```

```

; APPLICANT: Hansson, Lennart

```

```

; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic

```

```

; TITLE OF INVENTION: Enzyme (SCCE)

```

```

; NUMBER OF SEQUENCES: 17

```

```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-557-146-12

Alignment Scores:
Pred. No.: 5,85e-115 Length: 225
Score: 1215.50 Matches: 223
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.11% Mismatches: 0
Query Match: 68.29% Indels: 1
Gaps: 2

US-09-905-083A-30 (1-969) x US-08-557-146-12 (1-225)

QY 103 ATTATGTGCGCCCATGTGCAAGAGGCTCCACCCATGCGAGGTGGCCCTGCTCAGT 162
Db 1 IleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuSer 20
QY 163 GGCATCAGTCTC--CACTGGGGAGCGCTCTGGTCAATGAGCGCTGGTCTCACTGCC 219
Db 21 GlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40
QY 220 GCCCACTGCAAGATGAATGAGTACACCGCTGCACCTGGCGCAGTGATACGCTGGCGCAGG 279
Db 41 AlaHisCysLysMetAsnGluTrpValHisLeuGlySerAspThrLeuGlyAspArg 60
QY 280 AGAGTCTAGAGATCAAGCGCTCGAAGTCATTCCGCCACCCCGGCTACTCCACACAGACC 339
Db 61 ArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyrSerThrGlnThr 80
QY 340 CATGTTATGACCTCATGCTGTGAGCTCAATAGCCAGCGCCAGCTGTCTCATCGTG 399
Db 81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100
QY 400 AAGAAAGTCAGCTCCCTCCCGCTGCGAACCCCTCGAACCCCTGACTGTCTCCGCGC 459
Db 101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120
QY 460 TGGGCACTACACGAGCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGAATGC 519
Db 121 TrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal 140
QY 520 AAGCTCATCTCCCCCAGAGCTGACGAAAGTTTACAGAGGACTTACTGGAAAATTCATG 579
Db 141 LysLeuIleSerProGlnAspCysThrGluValTyrLysAspLeuLeuGluAsnSerMet 160

CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12

Alignment Scores:
Pred. No.: 5,85e-115 Length: 225
Score: 1215.50 Matches: 223
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.11% Mismatches: 0
Query Match: 68.29% Indels: 1
Gaps: 2

US-09-905-083A-30 (1-969) x US-09-154-344-12 (1-225)

QY 580 CTGTGCGTGGCATCCCGCATCCAGAAAAACCGCTGCAATGGTGACTCAGGGGACCG 639
Db 161 LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
QY 640 TTGTGTGTCAGAGGTACCCCTGCAAGGTCTGGTGTCTCTGGGGAACTTTCCCTTGGGCCAA 699
Db 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln 200
QY 700 CCCAATGACCCAGGAGTCTACACCTCAAGTGTGCAAGTTTCAACAAGTGTGATAAATGACACC 759
Db 201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220
QY 760 ATGAAAAAGCATCGC 774
Db 221 MetLysLysHisArg 225

RESULT 14
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hanson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12

Alignment Scores:
Pred. No.: 5,85e-115 Length: 225
Score: 1215.50 Matches: 223
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.11% Mismatches: 0
Query Match: 68.29% Indels: 1
Gaps: 2

US-09-905-083A-30 (1-969) x US-09-154-344-12 (1-225)

QY 580 CTGTGCGTGGCATCCCGCATCCAGAAAAACCGCTGCAATGGTGACTCAGGGGACCG 639
Db 161 LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
QY 640 TTGTGTGTCAGAGGTACCCCTGCAAGGTCTGGTGTCTCTGGGGAACTTTCCCTTGGGCCAA 699
Db 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln 200
QY 700 CCCAATGACCCAGGAGTCTACACCTCAAGTGTGCAAGTTTCAACAAGTGTGATAAATGACACC 759
Db 201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220
QY 760 ATGAAAAAGCATCGC 774
Db 221 MetLysLysHisArg 225
```

```
QY 103 ATTATTGATGGGCCCCCATGTGTCAAGAGGCTCCACCCATGGCAGGTGGCCCTGTCTCAGT 162
Db 1 IIEIEAAsGlyValaProCysalaargGlySerHisProTrrpGlnValAlaLeuLeuSer 20
QY 163 GGCATACAGCTC---CACTGGCGAGCGCTCTGGTCAATGAGCGCTGGGTGTCTCACTGCC 219
Db 21 GYAENGLnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40
QY 220 GCCCACTGCAAGATCAATGATACACCGTGCACCTGGCGAGTACAGTGGCGGACAGG 279
Db 41 AlaHisCysLysSerMetAsnGluTrpValHisLeuGlySerAspThrLeuGlyAspArg 60
QY 280 AGAGCTCAGAGGATCAAGCGCTCGAAGTCAATTCGCGCACCCCGGTACTCCACACAGACC 339
Db 61 ArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyrSerThrGlnThr 80
QY 340 CATGTTAATGACTCATGCTCGTAAGCTCAATAGCCAGGCGAGCGTGTATCCATGATG 399
Db 81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100
QY 400 AAGAAGTCAAGGCTCCCTCCCGCTGCGAACCCTGGGAACCACTGTACTGTCTCCGCG 459
Db 101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120
QY 460 TGGGCGCACTTACCACGAGCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTC 519
Db 121 TrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal 140
QY 520 AAGCTCATCTCCCCCAGAGTGCACGAAGGTTTACAAGGACTTACTGGAAAATTTCCATG 579
Db 141 LysLeuLieserProGlnAspCysThrGluValTyrLysAspLeuLeuGluAsnSerMet 160
QY 580 CTGTGCGTGGCATCCCGCATCCAGAAAACCGCTGCAATGGTGTACTCAGGGGACCG 639
Db 161 LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
QY 640 TTGGTGTGCAGAGGTACCTCGAAGTCTGGTGTCTGGGGAACCTTCCCTTTGGCGCAA 699
Db 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln 200
QY 700 CCCAATGACCCAGAGTGTACACTCAAGTGTGCAAGTTTCAACCAAGTGGATTAATGACACC 759
Db 201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220
QY 760 ATGAATAAGCATCGC 774
Db 221 MetLysLysHisArg 225
```

## RESULT 15

```
US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
; OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar
; OTHER INFORMATION: domain in TAGD-12
US-09-261-416-7
```

Alignment Scores:

Pred. No.: 9.55e-76 Length: 154

```
Score: 829.00 Matches: 154
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 46.57% Indels: 0
DB: 3 Gaps: 0
US-09-905-083A-30 (1-969) x US-09-261-416-7 (1-154)
QY 178 TGGGAGGCGTCTGGTCAATGAGCGCTGGGTGTCTCACTGCGCGCCACTGCAGATGAAT 237
Db 1 CysGlyGlyValLeuValAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsn 20
QY 238 GAGTACACCGTGCACCTGGCGAGTGCATACGCTGGCGGACAGGAGCTCAGAGGATCAAG 297
Db 21 GluTyrThrValHisLeuGlySerAspThrLeuGlyAspArgAlaGlnArgIleLys 40
QY 298 GCCTCGAAGTCAATTCGCGCACCCCGGTACTCCACACAGCCCATGTTTAATGACCTCATG 357
Db 41 AlaSerLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAspLeuMet 60
QY 358 CTCGTGAAGCTCAATAGCCAGCGCCAGGCTGTCTCATCTCCATGGTGGAAGATCAGGCTGCC 417
Db 61 LeuValLysLeuAsnSerGlnAlaArgLeuSerMetValLysValArgLeuPro 80
QY 418 TCCGCTCGGAACCCCTCGGAACCACTGTACTGTCTCCGCTGGGCGCACTACACAGAGC 477
Db 81 SerArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThrSer 100
QY 478 CCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCGCCAG 537
Db 101 ProAspValThrPheProSerAspLeuMetCysValAspValLysLeuLieserProGln 120
QY 538 GACTGCGAAGAGTTTACAAGGACTTACTGGAAAATTTCCATGCTGTGGCTGGCATCCCC 597
Db 121 AspCysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIlePro 140
QY 598 GACTCCAAAGAAAACGCTCGCAATGGTGACTCAGGGGACCG 639
Db 141 AspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 154
```

Search completed: July 14, 2005, 17:34:56

Job time : 43.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: July 14, 2005, 17:32:06 ; Search time 169 Seconds  
(without alignments)  
4430.247 Million cell updates/sec

Title: US-09-905-083A-30  
Perfect score: 1780  
Sequence: 1 ggatttcgggctccatggc.....aagaacacacaaacccctcag 969

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 3452440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool/US09905083/runat\_14072005\_140356\_10721/app\_query.fasta\_1.1159  
-DB=Published Applications AA -QMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09905083 @CGN\_1\_1199 @runat\_14072005\_140356\_10721  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1364	76.6	253	9	US-09-888-615-98	Sequence 98, Appl
2	1364	76.6	253	9	US-09-764-762-3	Sequence 3, Appl
3	1364	76.6	253	14	US-10-071-214-2	Sequence 2, Appl
4	1364	76.6	253	14	US-10-071-214-48	Sequence 48, Appl
5	1364	76.6	253	14	US-10-264-283-90	Sequence 90, Appl
6	1364	76.6	253	15	US-10-295-027-498	Sequence 498, App
7	1364	76.6	253	15	US-10-173-999-48	Sequence 48, Appl
8	1364	76.6	253	16	US-10-408-765A-639	Sequence 639, App
9	1364	76.6	253	16	US-10-643-758A-95	Sequence 95, Appl
10	1364	76.6	253	17	US-10-948-518-95	Sequence 95, Appl
11	1364	76.6	253	17	US-10-868-490A-1	Sequence 1, Appl
12	1350	75.8	250	15	US-10-262-511-92	Sequence 92, Appl
13	1337	75.1	257	15	US-10-344-394-38	Sequence 38, Appl
14	1252	70.3	247	15	US-10-262-511-102	Sequence 102, App
15	1252	70.3	252	15	US-10-262-511-94	Sequence 94, Appl
16	1235	69.4	225	15	US-10-600-187-4	Sequence 4, Appl
17	1230	69.1	224	11	US-09-789-210-33	Sequence 33, Appl
18	1119	62.9	224	15	US-10-262-511-104	Sequence 104, App
19	1085.5	61.0	249	14	US-10-071-214-47	Sequence 47, Appl
20	1035.5	58.2	198	15	US-10-262-511-96	Sequence 96, Appl
21	1011.5	56.8	249	14	US-10-071-214-50	Sequence 50, Appl
22	996.5	56.0	243	14	US-10-071-214-46	Sequence 46, Appl
23	940.5	52.8	226	14	US-10-071-214-49	Sequence 49, Appl
24	937.5	52.7	181	15	US-10-262-511-98	Sequence 98, Appl
25	773	43.4	144	9	US-09-796-294-4	Sequence 4, Appl
26	773	43.4	144	14	US-10-461-787-4	Sequence 4, Appl
27	614	34.5	260	9	US-09-796-294-8	Sequence 8, Appl
28	614	34.5	260	14	US-10-461-787-8	Sequence 8, Appl
29	614	34.5	260	15	US-10-072-012-609	Sequence 609, App
30	612.5	34.4	293	9	US-09-739-907-82	Sequence 82, Appl
31	612.5	34.4	293	9	US-09-739-907-97	Sequence 97, Appl
32	612.5	34.4	293	9	US-09-989-722-309	Sequence 309, App
33	612.5	34.4	293	9	US-09-989-723-309	Sequence 309, App
34	612.5	34.4	293	9	US-09-989-279-309	Sequence 309, App
35	612.5	34.4	293	9	US-09-989-727-309	Sequence 309, App
36	612.5	34.4	293	9	US-09-989-731-309	Sequence 309, App
37	612.5	34.4	293	9	US-09-989-732-309	Sequence 309, App
38	612.5	34.4	293	9	US-09-991-073-309	Sequence 309, App
39	612.5	34.4	293	9	US-09-990-442-309	Sequence 309, App
40	612.5	34.4	293	9	US-09-991-163-309	Sequence 309, App
41	612.5	34.4	293	9	US-09-993-604-309	Sequence 309, App
42	612.5	34.4	293	9	US-09-990-456-309	Sequence 309, App
43	612.5	34.4	293	9	US-09-989-721-309	Sequence 309, App
44	612.5	34.4	293	9	US-09-992-598-309	Sequence 309, App
45	612.5	34.4	293	9	US-09-989-293A-309	Sequence 309, App

## ALIGNMENTS

RESULT 1  
US-09-888-615-98  
; Sequence 98 Application US/09888615  
; Patent No. US20020064856A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: CAENEPEEL, SEAN  
; APPLICANT: CHARYCZAK, GLEN  
; APPLICANT: MANNING, GERARD  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: NOVEL PROTEASES  
; FILE REFERENCE: 038602/1214  
; CURRENT APPLICATION NUMBER: US/09/888,615  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,047  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 150  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 98

```
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Alignment Scores:
Pred. No.: 1.16e-118 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
DB: 9 Gaps: 0

US-09-905-083A-30 (1-969) x US-09-888-615-98 (1-253)
QY 16 ATGGCAAGATCCCTTCTCTCCCTCGCAGATCTTACTGCTATCTTAGCTTGGAAACT 75
Db 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20
QY 76 GCAGGAGAAGAGCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC 135
Db 21 AlaGlyGluAlaGlnGlyAspLysIleleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATGGCAGGTGGCCCTGCTAGTGGCAATCGCTCCACTGGGAGGGCTCTGGTC 195
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCGCTGGGTGCTCACTGCGCCCACTGCAAGATGAATGATCACCGTGCACCTG 255
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpValHisLeu 80
QY 256 GCAGTGATGACGTGGGCGACAGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACACAGCCCATGTTAATGACCTCATGCTCGTGAAGCTCATAGC 375
Db 101 HisProGlyTyzSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCGAGCTGTCTATCCATGGTGAAGAAAGTCAAGGCTCGCCCTCCGCTGCGAACCCT 435
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCTACCAACAGAGCCAGATGTGACCTTTCCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCAAGAGTTTAC 555
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTy 180
QY 556 AAGGACTTTACTGGAAAAATTCATGCTGTGCGTGGCATCCCGCTCCCAAGAAAAAGCC 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TCCAATGCTGACTCAGGGGGACCGTTGTTGTGCAAGAGGTACCTCGCAAGGTCTGGTGTC 675
Db 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGGAATTTCCCTTGGGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyThrGlnValCysLys 240
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 2
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
```

```
; Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Alignment Scores:
Pred. No.: 1.16e-118 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
DB: 9 Gaps: 0

US-09-905-083A-30 (1-969) x US-09-764-762-3 (1-253)
QY 16 ATGGCAAGATCCCTTCTCTCCCTCGCAGATCTTACTGCTATCTTAGCTTGGAAACT 75
Db 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuAlaLeuGluThr 20
QY 76 GCAGGAGAAGAGCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC 135
Db 21 AlaGlyGluAlaGlnGlyAspLysIleleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATGGCAGGTGGCCCTGCTAGTGGCAATCAGCTCCACTGGGAGGGCTCTGGTC 195
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCGCTGGGTGCTCACTGCGCCCACTGCAAGATGAATGATCACCGTGCACCTG 255
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpValHisLeu 80
QY 256 GCAGTGATGACGTGGGCGACAGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACACAGCCCATGTTAATGACCTCATGCTCGTGAAGCTCATAGC 375
Db 101 HisProGlyTyzSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCGAGCTGTCTATCCATGGTGAAGAAAGTCAAGGCTCGCCCTCCGCTGCGAACCCT 435
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCTACCAACAGAGCCAGATGTGACCTTTCCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCAAGAGTTTAC 555
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTy 180
QY 556 AAGGACTTTACTGGAAAAATTCATGCTGTGCGTGGCATCCCGCTCCCAAGAAAAAGCC 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TCCAATGCTGACTCAGGGGGACCGTTGTTGTGCAAGAGGTACCTCGCAAGGTCTGGTGTC 675
Db 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGGAATTTCCCTTGGGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyThrGlnValCysLys 240
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGGCCAGCTGTCTATCCATGCTGAAGAAAGTCAGCGTCCCTCCGCTGGAGAACCCCT 435  
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCACAGATGTGACCTTTCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGCGTGGATGTCAGCTCATCTCCCCAGGACTCACGAAGGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAATTCATGTCTGGCTGGCATCCCGGCTCCAGAAACGCC 615  
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
QY 616 TGCATGTGACTCAGGGGGACCGTGTGTGTCAGAGGTACCTGCAAGGTCTGGTGTC 675  
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACCTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGCATCCG 774  
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

## RESULT 3

US-10-071-214-2  
; Sequence 2, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-214-2

Alignment Scores:  
Pred. No.: 1,16e-118 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
DB: 14 Gaps: 0

US-09-905-083a-30 (1-969) x US-10-071-214-2 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTGCCCCCTGCAGATCTTACTGCTATCTTAGCCTTGGAACT 75  
Db 1 MetAlaArgSerLeuLeuProLeuGlnLeuLeuSerLeuLeuAlaLeuGluThr 20  
QY 76 CGAGAGAAAGCCAGCGGTGACAAATATTGATGGCGCCCATGTGCAAGAGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCCATGCGAGGTGGCCCTGCTAGTGCATCAGTCCACTGCGGAGGCGTCTGTGTC 195

Db 41 HisProTrpGlnValAlaLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCCCTGGGTCTCTACTGCGCCCACTGCAAGATGAATGACTACACCGTGCACCTG 255  
Db 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80  
QY 256 GGCAGTGTATCGCTGGGGACAGAGAGCTCAGAGGATCAAGGCTCAGAGTCAATTCGCG 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120  
QY 376 CAGGCCAGCTGTCTATCCATGCTGAAGAAAGTCAGGCTGCCCTCCGCTGGAGAACCCCT 435  
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCACAGATGTGACCTTTCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGCGTGGATGTCAGCTCATCTCCCCAGGACTCACGAAGGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAATTCATGTCTGGCTGGCATCCCGGCTCCAGAAACGCC 615  
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
QY 616 TGCATGTGACTCAGGGGGACCGTGTGTGTCAGAGGTACCTGCAAGGTCTGGTGTC 675  
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACCTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGCATCCG 774  
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

## RESULT 4

US-10-071-214-48  
; Sequence 48, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from  
; OTHER INFORMATION: homo sapiens.  
US-10-071-214-48

Alignment Scores:  
Pred. No.: 1,16e-118 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0



```
Query Match: 76.63% Indels: 0
DB: 14 Gaps: 0
US-09-905-083A-30 (1-969) x US-10-071-214-48 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTCTCCCTGCAGATCTTACTGCTATCTTAGCCTTGGAACCT 75
DB 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuThr 20
QY 76 GCAGGAGAGAGAGCCAGGGTGACAAAGATTATTGATGGCGCCCAATGTCGAAGAGCTCC 135
DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATGGCAGGTGGCCCTGCTAGTGGCAATCAGCTCCAGTCCAGTGGGAGGCGTCTGTGTC 195
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCCTGGTCTCACTGCGCCCAATGAGTGAATGAGTCAACCGTGACCTG 255
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80
QY 256 GGCAGTGTATCGCTGGGCGCAGGAGACTCAGAGGATCAAGGCTCGAAGTCTTCGCC 315
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACAGACCATGTTAATGACCTCATGCTCGTGAAGTCAATAGC 375
DB 101 HisProGlyTrpSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCGAGGCTCATCTCCGCTGGGCGCAGGAGTCAAGGCTCGGCTCGCAACCCCT 435
DB 121 GlnAlaArgLeuSerMetValLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGCTGGGCGCAGGAGTCAAGGCTCGGCTCGCAACCCCT 495
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATGTGCTGGATGTCAGCTCATCTCCCCCAGAGCTGCACGAAGTTTAC 555
DB 161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180
QY 556 AAGGACTTACTGGAATAATCCATGCTGCGCTGGCATCCCGCTCCCAAGAAACGCC 615
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TGCATGTGTACTGGAATAATCCATGCTGCTGGCTGGCATCCCGCTCCCAAGAAACGCC 675
DB 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGAACTTTCCCTTGGGAAATTCATGCTGTGCTGGCTGGCATCCCGCTCCCAAGAAACGCC 735
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGCATCGC 774
DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 5
US-10-264-283-90
; Sequence 90, Application US/10264283
; Publication No. US2003014494A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.590
; CURRENT APPLICATION NUMBER: US/10/264,283
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 90
; LENGTH: 253

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-283-90

Alignment Scores:
Pred. No.: 1,16e-118 Length: 253
Score: 1364.00 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
DB: 14 Gaps: 0
US-09-905-083A-30 (1-969) x US-10-264-283-90 (1-253)

QY 16 ATGGCAAGATCCCTTCTCTCTCCCTGCAGATCTTACTGCTATCTTAGCCTTGGAACCT 75
DB 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuThr 20
QY 76 GCAGGAGAGAGAGCCAGGGTGACAAAGATTATTGATGGCGCCCAATGTCGAAGAGCTCC 135
DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40
QY 136 CACCATGGCAGGTGGCCCTGCTAGTGGCAATCAGCTCCAGTCCAGTGGGAGGCGTCTGTGTC 195
DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60
QY 196 AATGAGCCTGGTCTCACTGCGCCCAATGAGTGAATGAGTCAACCGTGACCTG 255
DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTrpThrValHisLeu 80
QY 256 GGCAGTGTATCGCTGGGCGCAGGAGACTCAGAGGATCAAGGCTCGAAGTCTTCGCC 315
DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100
QY 316 CACCCGGCTACTCCACAGACCATGTTAATGACCTCATGCTCGTGAAGTCAATAGC 375
DB 101 HisProGlyTrpSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
QY 376 CAGGCGAGGCTCATCTCCGCTGGGCGCAGGAGTCAAGGCTCGGCTCGCAACCCCT 435
DB 121 GlnAlaArgLeuSerMetValLysValArgLeuProSerArgCysGluProPro 140
QY 436 GGAACCACTGTACTGTCTCCGCTGGGCGCAGGAGTCAAGGCTCGGCTCGCAACCCCT 495
DB 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
QY 496 TCTGACCTCATGTGCTGGATGTCAGCTCATCTCCCCCAGAGCTGCACGAAGTTTAC 555
DB 161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180
QY 556 AAGGACTTACTGGAATAATCCATGCTGCGCTGGCATCCCGCTCCCAAGAAACGCC 615
DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
QY 616 TGCATGTGTACTGGAATAATTCATGCTGTGCTGGCTGGCATCCCGCTCCCAAGAAACGCC 675
DB 201 CysAsnGlyAspSerGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220
QY 676 TGGGAACTTTCCCTTGGGAAATTCATGCTGTGCTGGCTGGCATCCCGCTCCCAAGAAACGCC 735
DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240
QY 736 TTCACCAAGTGGATAAATGACACCATGAAAGCATCGC 774
DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 6
US-10-295-027-498
; Sequence 498, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
```

Db	101	HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer	120
Qy	376	CAGGCACGGCTGTCATCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCT	435
Db	121	GlnAlaArgLeuSerSerMetValLysValArgLeuProSerArgCysGluProPro	140
Qy	436	GGAACCACTGACTGCTCCCGGCTGGGCACTACCAAGAGCCAGATGTGACCTTTCC	495
Db	141	GlyThrThrCysThrValSerGlyTyrGlyThrThrThrSerProAspValThrPhePro	160
Qy	496	TCTGACCTCATGTGGTGGATGTCAAGTCTCATCTCCCCCAGGACTGCACGAAGGTTTAC	555
Db	161	SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr	180
Qy	556	AAGGACTTACTGGAATTCATGCTGTGCGCTGGCACTCCCGACTCCCAAGAAAAACGCC	615
Db	181	LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla	200
Qy	616	TGCAATGCTGACTCAGGGGACCGTGTGTGTGCAGAGTACCTGCAAGGTCTGGTGTCC	675
Db	201	CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer	230
Qy	676	TGGGGAACCTTCCCTTGGGGCAACCCAAATGACCAGGAGTCTACTCAAGTGTGCAAG	735
Db	221	TyrGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys	240
Qy	736	TTCAACCAAGTCGATAAATGACCACTGAAAGCAATCGC	774
Db	241	PheThrLysTyrIleAsnAspThrMetLysLysHisArg	253

RESULT 7

US-10-173-999-48

; Sequence 48, Application US/10173999

; Publication No. US2004000563A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions

; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian

; TITLE OF INVENTION: Cancer

; FILE REFERENCE: 018501-00242005

; CURRENT APPLICATION NUMBER: US/10/173,999

; CURRENT FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: US 60/299,234

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: US 60/315,287

; PRIOR FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/372,246

; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 163

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 48

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-173-999-48

Alignment Scores:	
Pred. No.:	1,16e-118
Score:	1364.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	76.63%
DB:	15
	Length: 253
	Matches: 253
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-09-905-083A-30 (1-969) x US-10-173-999-48 (1-253)

Qy	16	ATGCAAGATCCCTTCTCCTCCCTCGAGATCTTACTGCTTACTGCTTACGCTTGAAACT	75
----	----	------------------------------------------------------------	----

Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGTCACAGATATTATGATGGCGCCCATGTGCAAGAGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGCGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCTGTGTC 195  
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAenGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCCTGGTCTCAGTGGCCGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCTGTGTC 255  
Db 61 AenGluArgTrpValLeuThrAlaAlaHisCysLysMetAenGluThrValHisLeu 80  
QY 256 GCGAGTATACGCTCGCGGACAGGAGCTCAGAGGATCAAGGCTCGCAAGTCATTCGCGC 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCATTGTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAenAspLeuMetLeuValLysLeuAenSer 120  
QY 376 CAGCGCCAGCTGATCCTCAGTGGTGAAGAGTCAAGCTGGCTGCGCTCGCAAGCCCTC 435  
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GAAACCACTGTACTGTCTCGCGCTGGGCGCACTACACAGAGCCAGATGTGACCTTTCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
QY 496 TGTGACCTCATGTGCGTGGATGTAAGCTCATCTCCCGCCAGGACTGCAAGAGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGCTTACTGGAATTCATGTGTGCGTGGCATCCCGCTCCAGGATCAAGAAACGCC 615  
Db 181 LysAspLeuLeuGluAenSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
QY 616 TGCATGTGTACTCAGGGGACCGTGTGTGTGTCAGAGGTACCTGCAAGGTCTGTGTC 675  
Db 201 CysAenGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAenAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATTAATGACACCATGAAGAAGCATCGC 774  
Db 241 PheThrLysTrpIleAenAspThrMetLysLysHisArg 253

## RESULT 8

US-10-408-765A-639  
; Sequence 639, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088, 465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 639  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-10-408-765A-639

Alignment Scores:  
Pred. No.: 1,16e-118 Length: 253  
Score: 1364.00 Matches: 253  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 76.63% Indels: 0  
Gaps: 16

## US-09-905-083A-30 (1-969) x US-10-408-765A-639 (1-253)

QY 16 ATGCAAGATCCCTTCTCTCTCCCTGCAGATCTTACTGTCTATCTTAGCCTTGAAACT 75  
Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20  
QY 76 GCAGGAGAGAGCCAGGTCACAGATTTATTGATGGCGCCCATGTGCAAGAGGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGCGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCTGTGTC 195  
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAenGlnLeuHisCysGlyValLeuVal 60  
QY 196 AATGAGCCTGGTCTCAGTGGCCGCTCAGTGGCAATGAAATGATGATACACCGTGACCTG 255  
Db 61 AenGluArgTrpValLeuThrAlaAlaHisCysLysMetAenGluThrValHisLeu 80  
QY 256 GCGAGTATACGCTGGGCGACAGGAGCTCAGAGGATCAAGGCTCGAAGCTCATTCGCGC 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGCTACTCCACACAGACCATTGTAATGACCTCATGCTCGTGAAGCTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAenAspLeuMetLeuValLysLeuAenSer 120  
QY 376 CAGCGCCAGCTGTATCATCGTGAAGAAATCAGGCTGCGCTCCCGCTGCGAAGCCCTC 435  
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCGCGTGGGCGCACTACACAGGAGCCAGAGTGTGACCTTTCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
QY 496 TGTGACCTCATGTGCGTGGATGTAAGCTCATCTCCCGCCAGGACTGCAAGAGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAATTCATGTGTGCGTGGCATCCCGCTCCAGGATCAAGAAACGCC 615  
Db 181 LysAspLeuLeuGluAenSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200  
QY 616 TGCATGTGTACTCAGGGGACCGTGTGTGTGTCAGAGGTACCTGCAAGGTCTGTGTC 675  
Db 201 CysAenGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGAACTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAenAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGGATTAATGACACCATGAAGAAGCATCGC 774  
Db 241 PheThrLysTrpIleAenAspThrMetLysLysHisArg 253

## RESULT 9

US-10-643-795A-95  
; Sequence 95, Application US/10643795A  
; Publication No. US20040241703A1  
; GENERAL INFORMATION:  
; APPLICANT: FREDERIC J. DESAUVAGE  
; APPLICANT: GRETCHEN PRANTZ  
; APPLICANT: KENNETH J. HILLAN  
; APPLICANT: PAUL POLAKIS





```

; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 92
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-92

Alignment Scores:
Pred. No.: 2,376-117 Length: 250
Score: 1350.00 Matches: 250
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 75.84% Indels: 0
DB: 15 Gaps: 0

US-09-905-083A-30 (1-969) x US-10-262-511-92 (1-250)

QY 25 TCCCTCTCTCCGCGCCCTGCAGATCTTACTGCTATCTTACTGCTTGGAACTGCAGAGAA 84
DB 1 SerLeuLeuLeuProLeuGlnLeuLeuLeuSerLeuLeuLeuLeuLeuLeuThraLeuGlu 20
QY 85 GAAGCCAGGGTGACAAGATTATTGATCGCGCCCATGTGCAAGAGGCTCCCAACCCATGG 144
DB 21 GluAlaGlnGlyAspIleIleAspGlyAlaProCysAlaArgGlySerHisProTrp 40
QY 145 CAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTCGGAGGCGTCTCGTCAATGAGCGC 204
DB 41 GlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArg 60
QY 205 TGGGTGCTCACTGCGCGCCCACTGCAAGATGAATGATGATACCGGTGCACCTGGGCGATGAT 264

61 TrpValLeuThraAlaAlaHisCysLeuMetAsnGluTyrThrValHisLeuGlySerAsp 80
265 ACGCTGGGCGCAGAGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCCGCCACCCCGGC 324
81 ThrLeuGlyAspArgAlaGlnArgIleLeuAlaSerLeuSerPheArgHisProGly 100
325 TACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCAGG 384
101 TyrSerThrGlnThrHisValAsnAspLeuMetLeuValLeuLeuAsnSerGlnAlaArg 120
385 CTGTCAATCCATGGTGAAGAAAGTCAGGCTGCGCTCCCGCTCGGAACCCCTCGGAACACC 444
121 LeuSerSerMetValLeuLeuValArgLeuProSerArgCysGluProProGlyThrThr 140
445 TGTACTGTCTCCGCTGGGCGCACTACACAGAGCCAGATGACCTTTCCCTCTGACCTC 504
141 CysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeu 160
505 ATGTGCGTGGATGTCAGAGCTCATCTCCCGCCAGAGCTGCACGAAGGTTTCAAGGACTTA 564
161 MetCysValAspValLeuLeuLeuSerProGlnAspCysThrLysValTyrLysAspLeu 180
565 CTGGAATAATCCATGCTGTGGCTGGCATCCCGCATCCCGACTCCAAGAAAAACGCTGCAATGGT 624
181 LeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGly 200
625 GACTCAGGGGACCGTTGGTGTGCAGAGGTACCTGCAAGGCTCTGGTGTCTGGGGAAC 684
201 AspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThr 220
685 TTCCCTTCGCGGCAACCCCAATGACCCAGAGGTCTACACTCAAGTGTGCAAGTTCACCAAG 744
221 PheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLys 240
745 TGGATAAATGACACCATGCAAAAAGCATCGC 774
241 TrpIleAsnAspThrMetLysLysHisArg 250

RESULT 13
US-10-344-394-38
; Sequence 38, Application US/10344394
; Publication No: US20040058342A1
; GENERAL INFORMATION:
; APPLICANT: Yousef, George M.
; TITLE OF INVENTION: NOVEL KALLIKREIN GENE
; FILE REFERENCE: 11757.51USWO
; CURRENT APPLICATION NUMBER: US/10/344,394
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: PCT/CA01/01141
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/224,853
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-344-394-38

Alignment Scores:
Pred. No.: 3,94e-116 Length: 257
Score: 1337.00 Matches: 252
Percent Similarity: 97.67% Conservative: 0
Best Local Similarity: 97.67% Mismatches: 0
Query Match: 75.11% Indels: 6
DB: 15 Gaps: 2

US-09-905-083A-30 (1-969) x US-10-344-394-38 (1-257)

QY 16 ATGCAAGATCCCTTCTCGCCCTGCAGATCTTACTGCTATCTTACTGCTTGGAACT 75
```

Db 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuLeuLeuSerLeuAlaLeuGluThr 20  
Qy 76 GCAGCAGAGAGCCAGCGGTGACAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC 135  
Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
Qy 136 CACCATTGCGCAGGTGGCCCTGTCTAGTGGCAATCAGCTCCACTGC----- 180  
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysHisSerCysCysGlu 60  
Qy 181 GGAGCGTCTGGTCAATAGCGCTGGGTGCTCACTGGCGCCACTGCAAGATGAATGAG 240  
Db 61 GlyGlyValLeuValAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGlu 80  
Qy 241 TACACCGTGCACTGGCGCAGTGATACGCTGGCGGACGAGAGCTCAGAGATCAAGGCC 300  
Db 81 TyrThrValHisLeuGlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAla 100  
Qy 301 TCGAAGTCAATTCGCGCACCCTGGCTACTCCACAGACCCCATGTTTAATGACCTCATGCTC 360  
Db 101 SerLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeu 120  
Qy 361 GTGAAGCTCAATAGCAGCGCAGCTGTCTCATCTGTTGAAGAAAGTCAGGCTGCCCTCC 420  
Db 121 ValLysLeuAsnSerGlnAlaArgLeuSerMetValLysLysValArgLeuProSer 140  
Qy 421 CGTCGCAACCCCTGGGAACCACTGTACTGTCTCCGGCTGGCGCACTACCACGAGCCCA 480  
Db 141 ArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerPro 160  
Qy 481 GATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCGCCAGGAC 540  
Db 161 AspValThrPhePro---AspLeuMetCysValAspValLysLeuLysSerProGlnAsp 179  
Qy 541 TGCAGAGGTTTACAAGACTTACTGGAATAATTCATGCTGTGGCGCTGGCATCCCGAC 600  
Db 180 CysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAsp 199  
Qy 601 TCCAGAAAGAGCGCTGCAATGGTGACTCAGCGGAGCGTGTGGTGCGAGAGGTACCGTG 660  
Db 200 SerLysLysAsnAlaCysAsnGlyAspSerGlyLysProLeuValCysArgGlyThrLeu 219  
Qy 661 CAAGTCTGTGTCTCTGGGGAACCTTTCCCTTTCGCGCCCAACCAATGACCCAGGAGTCTAC 720  
Db 220 GlnGlyLeuValSerTrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyr 239  
Qy 721 ACTCAAGTGTCAAGTTCACCAAGTGGGATAATGACACCATGACCCCAAGCATCGC 774  
Db 240 ThrGlnValCysLysPheThrLysTrpIleAsnAspThrMetLysLysHisArg 257

RESULT 14

US-10-262-511-102  
; Sequence 102, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerkhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei

US-09-905-083A-30 (1-969) x US-10-262-511-102 (1-247)  
Qy 88 GCCCAGGGTGCACAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCACCCATGGCAG 147  
Db 11 AlaArgGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGln 30  
Qy 148 GTGGCCCTGTCTAGTGGCAATCAGCTCACTGGCGAGCGCTCTGGTCAATGAGGCTGG 207  
Db 31 ValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrp 50  
Qy 208 GTGCTCACTGCGCGCCACTGCAAGATGAATGAGTACACCGTGACCTGGGCGAGTGATACG 267  
Db 51 ValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThr 70  
Qy 268 CTGGCGCAGAGGAGGAGCTCAGAGGATCAAGGCTTCGAAGTCAATTCGCGCCACCCCGGCTAC 327  
Db 71 LeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyr 90  
Qy 328 TCCACACAGACCCCATGTTAATGACCTCATGCTGTGAAGCTCAATAGCCAGCCAGCGCTG 387

Alignment Scores:  
Pred. No.: 3.55e-108 Length: 247  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 70.34% Indels: 0  
DB: 15 Gaps: 0

US-10-262-511-102  
; ORGANISM: Homo sapiens  
; TYPE: PRT  
; LENGTH: 247  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 102  
; NUMBER OF SEQ ID NOS: 439  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; SOFTWARE: CuraSeqList version 0.1



Db 91 SerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeu 110  
Qy 388 TCATCATGCTGGAAGAACTCAGGCTGCCCTCCGCTCGAACCCTCGGACCACTGT 447  
Db 111 SerSerMetValLysValArgLeuProSerArgCysGluProGlyThrThrCys 130  
Qy 448 ACTGCTCCGGCTGGGGCACTACACGAGCCAGATGTGACCTTTCCCTCTGACCTCATG 507  
Db 131 ThrValSerGlyTyrGlyThrThrSerProAspValThrPheProSerAspLeuMet 150  
Qy 508 TCGTGGATGTCAAGCTCATCTCCCGCAGAGCTCAGCAAGGTTTCAAGACTTACTG 567  
Db 151 CysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyrLysAspLeuLeu 170  
Qy 568 GAAATTCATGCTGCTGGCTGGCATCCCGACTCCAGAAACCGCTGCAATGGTGCAC 627  
Db 171 GluAsnSerMetLeuCysAlaGlyLeuProAspSerLysLysAsnAlaCysAsnGlyAsp 190  
Qy 628 TCAGGGGACCGTTCGTGTGCAGAGGTACCTCGTCAAGGTCTGCTGCTGGGAACTTTC 687  
Db 191 SerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTyrGlyThrPhe 210  
Qy 688 CTTTCGGGCAACCACTCACTCCCGCAGAGCTTACACTCAAGTGTGCAAGTTCACCAAGTGG 747  
Db 211 ProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrp 230  
Qy 748 ATAATGACACATCAAAAGCATCGC 774  
Db 231 IleAsnAspThrMetLysLysHisArg 239

## RESULT 15

US-10-262-511-94  
; Sequence 94, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Saasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zehrusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917

Alignment Scores:  
Pred. No.: 3,57e-108 Length: 252  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 70.34% Indels: 0  
DB: 15 Gaps: 0

US-09-905-083a-30 (1-969) x US-10-262-511-94 (1-252)  
Qy 88 GCCCAGGTGACAAGATTATTGATGGCGCCCATGTGTCAAGAGGCTCCACCCATGGCAG 147  
Db 24 AlaArgGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGln 43  
Qy 148 GTGGCCCTGCTCAGTGGCAATCAGCTCCAGCTGGCGAGCGCTCTGGTCAATGAGGCGTGG 207  
Db 44 ValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrp 63  
Qy 208 GTGTCTACTGCGGCCCTCAAGATGAATCAGTACACCGTGACCTGGGAGTGATACG 267  
Db 64 ValLeuThrAlaAlaHisCysLysMetAsnGluThrThrValHisLeuGlySerAspThr 83  
Qy 268 CTGGCGCAGCAGGAGCTCAGAGGATCAAGCGCTCGAAGTCAATTCGCCACCCCGGCTAC 327  
Db 84 LeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyr 103  
Qy 328 TCCACACAGACCCATGTTATGACTCATGTCTGTGAAGCTCAATAGCAGCCAGGCTG 387  
Db 104 SerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeu 123  
Qy 388 TCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTCGCAACCCCTCGGAACCACTGT 447  
Db 124 SerSerMetValLysLysValArgLeuProSerArgCysGluProGlyThrThrCys 143  
Qy 448 ACTGTCTCCGGCTGGGCACTTACCAGCCAGCTGAGTGTGACCTTTCCCTCTGACCTCATG 507  
Db 144 ThrValSerGlyTyrGlyThrThrSerProAspValThrPheProSerAspLeuMet 163  
Qy 508 TGCCTGATGTCAAGCTCATCTCCCGCAGAGCTGCGAAGGTTTCAAGACTTACTG 567  
Db 164 CysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyrLysAspLeuLeu 183  
Qy 568 GAAATTCATGCTGGCTGGCATCCCGACTCCAGAAACCGCTGCAATGGTGCAC 627  
Db 184 GluAsnSerMetLeuCysAlaGlyLeuProAspSerLysLysAsnAlaCysAsnGlyAsp 203  
Qy 628 TCAGGGGACCGTTCGTGTGCAGAGGTACCTCCCAAGGCTCTGGTGTCTGGGAACTTTC 687  
Db 204 SerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTyrGlyThrPhe 223

QY	688	CCTTCGGCCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGG	747
Db	224	ProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrp	243
QY	748	ATAATGACACCATGAAAAAGCATCGC	774
Db	244	IleAsnAspThrMetLysLysHisArg	252

Search completed: July 14, 2005, 17:57:24  
Job time : 178 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 14, 2005, 17:15:20 ; Search time 45.5 Seconds

(without alignments)

4098.200 Million cell updates/sec

Title: US-09-905-083A-30

Perfect score: 1780

Sequence: 1 ggatttcgggctccatggc.....aagaacacacacccctcag 969

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09905083/runat\_14072005\_140353\_10622/app\_query.fasta\_1.1159  
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=expct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09905083 @CGN\_1\_1\_63 @runat\_14072005\_140353\_10622 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1364	76.6	253	2 A53968	serine proteinase
2	614	34.5	260	2 I56559	neuropain - mouse
3	582	32.7	261	2 A29745	tissue kallikrein
4	568	31.9	261	2 S01971	tissue kallikrein
5	563	31.6	261	1 NGMSG	7S nerve growth fa
6	563	31.6	261	2 A34079	tissue kallikrein
7	561	31.5	246	1 TRRT1	trypsin (EC 3.4.21
8	552	31.0	261	1 KQMS1	tissue kallikrein
9	552	31.0	261	2 A31136	tissue kallikrein
10	551	31.0	260	2 A25606	tissue kallikrein
11	547.5	30.8	261	2 A37938	tissue kallikrein
12	541.5	30.4	246	1 TRRT2	trypsin (EC 3.4.21
13	536	30.1	248	2 S55067	trypsin (EC 3.4.21
14	535	30.1	229	1 TRB0TR	trypsin (EC 3.4.21

15	535	30.1	265	1 KQTP	tissue kallikrein
16	534	30.0	261	1 TRSM5	tissue kallikrein
17	533.5	30.0	244	2 A44284	tissue kallikrein
18	533.5	30.0	247	2 A27547	trypsin (EC 3.4.21
19	533	29.9	246	2 B25528	trypsin (EC 3.4.21
20	533	29.9	259	1 KQRTTN	tonin (EC 3.4.21.-
21	533	29.9	259	2 A29746	tissue kallikrein
22	533	29.9	259	2 B31136	tissue kallikrein
23	532.5	29.9	231	2 S31778	trypsin (EC 3.4.21
24	532	29.9	242	2 S31776	trypsin (EC 3.4.21
25	532	29.9	242	2 S31775	trypsin (EC 3.4.21
26	531	29.8	261	2 A41020	tissue kallikrein
27	527.5	29.6	247	1 TRDG	trypsin (EC 3.4.21
28	527	29.6	261	2 JF0236	tissue kallikrein
29	526.5	29.6	242	2 S49489	trypsin (EC 3.4.21
30	526.5	29.6	247	2 S13813	trypsin (EC 3.4.21
31	524	29.4	231	1 TRPGTR	trypsin (EC 3.4.21
32	521	29.3	261	2 A24378	tissue kallikrein
33	519	29.2	261	1 EGMSB	tissue kallikrein
34	518.5	29.1	243	2 A35871	trypsin (EC 3.4.21
35	518.5	29.1	262	1 KQHU	tissue kallikrein
36	516.5	29.0	256	1 NGMSA	7S nerve growth fa
37	516	29.0	257	2 S33772	tissue kallikrein
38	514	28.9	232	1 KQPG	tissue kallikrein
39	513	28.8	261	2 A29586	tissue kallikrein
40	511	28.7	238	2 S31779	trypsin (EC 3.4.21
41	510.5	28.7	247	1 A25852	trypsin (EC 3.4.21
42	509	28.6	263	2 S15686	tissue kallikrein
43	508	28.5	246	1 TRDGC	tissue kallikrein
44	506	28.4	261	2 S45303	tissue kallikrein
45	503	28.3	248	2 S55066	trypsin (EC 3.4.21

#### ALIGNMENTS

##### RESULT 1

A53968

N;Alternate names: stratum corneum chymotryptic enzyme

C;Species: Homo sapiens (man)

C;Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C;Accession: A53968

R;Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A;Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A;Reference number: A53968; MUID:94308225; PMID:8034709

A;Status: preliminary

A;Accession: A53968

A;Residues: 1-253 <HAN>

A;Cross-references: UNIPROT:P49862; GB:L33404; MID:9521214; PIDN:AAC37551.1; PID:9532504

C;Genetics:

A;Gene: GDB:PRSS6; SCCE

A;Cross-references: GDB:377730

A;Map position: 7q35-7q35

C;Superfamily: trypsin; trypsin homology

F;30-245/Domain: trypsin homology <TRY>

##### Alignment Scores:

Pred. No.:	3.28e-109	Length:	253
Score:	1364.00	Matches:	253
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	76.63%	Indels:	0
DB:	2	Gaps:	0

US-09-905-083A-30 (1-969) x A53968 (1-253)

Oy	16	ATGCCAGATCCCTTCCTCCCTGCGAGATCTTACTGCTATCTTACCTTGGAACT	75
Db	1	MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr	20
Oy	76	GCAGAGAGAGAGAGCCAGGGTGACAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC	135

Db 21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40  
QY 136 CACCATGCGAGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGCTCTGTC 195  
Db 41 HisProTrpGlnValAlaLeuLeuSerGlyAenGlnLeuHisCysGlyGlyValLeuVal 60  
QY 196 AATGAGCGTGGGTGCTCACTCCGCCCACTCAAGAGTAATGAGTACACCGTGCACCTG 255  
Db 61 AenGluAArgTrpValLeuThraAlaHisCysLysMetAenGluTrpThrValHisLeu 80  
QY 256 GCGAGTGATACCTGGCGCAGAGAGCTCAGAGATCAAGGCTCGAAGTCATTCGCG 315  
Db 81 GlySerAspThrLeuGlyAspArgAlaGlnAArgIleLysAlaSerLysSerPheArg 100  
QY 316 CACCCCGGCTACTCCACAGACCCATGTTAATGACCTCATCTCGTGAAGTCAATAGC 375  
Db 101 HisProGlyTyrSerThrGlnThrHisValAenAspLeuMetLeuValLysLeuAenSer 120  
QY 376 CAGGCGAGGCTGCTCATCCATGCTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCT 435  
Db 121 GlnAlaArgLeuSerSerMetValLysLysValAArgLeuProSerArgCysGluProPro 140  
QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCAGAGTGTGACCTTTCCC 495  
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160  
QY 496 TCTGACCTCATGTGGTGGTGAAGTCAAGCTCATCTCCCGCCAGGAGTGCACGAAGTTTAC 555  
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180  
QY 556 AAGGACTTACTGGAATTTCCATGTGTGCGTGGCATCCCCGACTCCAAGAAAAACGCC 615  
Db 181 LysAspLeuLeuGluAenSerMetLeuCysAlaGlyIleProAspSerLysLysAenAla 200  
QY 616 TCATATGTGACTCAGAGGGGACCGTGTGTGTCAGAGGTACCTCGAAGGTGTGTGTGCC 675  
Db 201 CysAenGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220  
QY 676 TGGGGAACCTTCCCTTGGCGGCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 221 TrpGlyThrPheProCysGlyGlnProAenAspProGlyValTyrThrGlnValCysLys 240  
QY 736 TTCACCAAGTGTGATAAATGACACCATGAAAGCATCGC 774  
Db 241 PheThrLysTrpIleAenAspThrMetLysLysHisArg 253  
RESULT 2  
156559  
neurospisin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: 156559  
R:Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Nishih  
J. Neurosci. 15, 5088-5097, 1995  
A:Title: Expression and activity-dependent changes of a novel limbic-serine protease gen  
A:Reference number: 156559; MUID:95348817; PMID:7623137  
A:Accession: 156559  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-260 <RES>  
A:Cross-references: UNIPROT:Q61955; GB:D30785; NID:G1648847; PIDN:BAA06451.1; PID:G10200  
C:Superfamily: trypsin; trypsin homology  
F:33-252/Domain: trypsin homology <TRY>  
Alignment Scores:  
Pred. No.: 1.05e-44 Length: 260  
Score: 614.00 Matches: 114  
Percent Similarity: 61.94% Conservative: 39  
Best Local Similarity: 46.15% Mismatches: 88  
Query Match: 34.43% Indels: 6  
Gaps: 2  
DB:

US-09-905-083A-30 (1-969) x 156559 (1-260)  
QY 37 CCCCTGCAGATCTTACTGCTATCTTAGCCCTTGAAACTGCAGGAGAAGAGCCAGGCT 96  
Db 11 ProTrpIleLeuLeuLeuPheMetGlyAlaTrpAlaGlyLeuThrArgAlaGlnGly 30  
QY 97 GACAAGATTATTGATGGCGCCCCCATGTGCAAGAGGCTCCCAACCATGCGAGGTGGCCCTG 156  
Db 31 SerLysIleLeuGluGlyArgGluCysIleProHisSerGlnProTrpGlnAlaAlaLeu 50  
QY 157 CTCAGTGGCAATCAGCTCCACTGCGGAGGCGCTCTGCTCAATGAGCGTGGTGTCTCACT 216  
Db 51 PheGlnGlyGluArgLeuIleCysGlyGlyValLeuValGlyAspArgTrpValLeuThr 70  
QY 217 GCGCGCCACTCAAGATGAATGAGTACACCTGCACCTGGCGCACTGGGCACTGATACGCTG 270  
Db 71 AlaAlaHisCysLysLysGlnLysTrpSerValArgLeuGlyAspHisSerLeuGlnSer 90  
QY 271 GCGCAGCAGGAGAGCTCAGAGGATCAAGGCTTCGAAGTCAATTCGCCCAACCCCGGTACTCC 330  
Db 91 ArgAspGlnProGluGlnGluIleGlnValAlaGlnSerIleGlnHisProCysTyrAsn 110  
QY 331 ACA-----CAGACCCATGTTAATGACCTCATCTCGTGAAGCTCAATAGCAGGCC 381  
Db 111 AenSerAsnProGluAspHisSerHisAspIleMetLeuIleArgLeuGlnAenSerAla 130  
QY 382 AGGTGTCATCATGTGTGAAGAAAGTCAAGGCTGCGCTCCCGCTCGAAACCCCTGGAAACC 441  
Db 131 AenLeuGlyAspLysValLysProValGlnLeuAlaAenLeuCysProLysValGlyGln 150  
QY 442 ACCTGTACTGTCTCCGCTGGGCACTACACAGAGCCAGATGTGACCTTTCCCTCTGCAC 501  
Db 151 LysCysIleIleSerGlyTrpGlyThrValThrSerProGlnGluAenPheProAenThr 170  
QY 502 CTCATGTCGTGGATGTCAAGCTCATCTCCCGCCAGAGCTGCACGAAGGTTTACAGGAC 561  
Db 171 LeuAenCysAlaGluValLysIleTyrSerGlnAenLysCysGluAArgAlaTyrProGly 190  
QY 562 TTACTGGAATAATTCATGCTGTGCGTGGCATCCCGACTCCCAAGAAAAACGCTGCAAT 621  
Db 191 LysIleThrGluGlyMetValCysAlaGly---SerSerAenGlyAlaAenThrCysGln 209  
QY 622 GGTCACTCAGGGGACCGCTGTGTGTGTCAGAGTACCCTCGAAGTCTGTGTCTCTGGGA 681  
Db 210 GlyAspSerGlyGlyProLeuValCysAspGlyMetLeuGlnGlyIleThrSerTrpGly 229  
QY 682 ACTTTCCTTCCGCGCCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTTACC 741  
Db 230 SerAspProCysGlyLysProGluLysProGlyValTyrThrLysIleCysArgTyrThr 249  
QY 742 AAGTGGATAAATGACACCATG 762  
Db 250 ThrTrpIleLysLysThrMet 256  
RESULT 3  
A29745  
tissue kallikrein (EC 3.4.21.35) mGK-9 precursor, submandibular - mouse  
N:Alternate names: glandular kallikrein mGK-9; major epidermal growth factor-binding pro  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
C:Accession: C29746; A29745; A27120; I70015  
R:Drinkwater, C.C.; Evans, B.A.; Richards, R.I.  
Biochemistry 26, 6750-6756, 1987  
A:Title: Mouse glandular kallikrein genes: identification and characterization of the ge  
A:Reference number: A90522; MUID:88107594; PMID:3322387  
A:Accession: C29746  
A:Molecule type: DNA  
A:Residues: 1-261 <DRI>  
A:Cross-references: UNIPROT:P15949; GB:M17985; NID:G193476; PIDN:AAA37681.1; PID:G387166  
A:Experimental source: strain BALB/c, salivary gland  
R:Blaber, M.; Isackson, P.J.; Bradshaw, R.A.  
Biochemistry 26, 6742-6749, 1987  
A:Title: A complete cDNA sequence for the major epidermal growth factor binding protein

A:Reference number: A29745; MUID:88107593; PMID:3322386  
A:Accession: A29745  
A:Molecule type: mRNA  
A:Residues: 1-261 <BLA>  
A:Cross-references: GB:M1762; NID:g192997; PIDN:AAA37541.1; PID:g309212  
R:Iackaon, P.J.; Silverman, R.E.; Blaber, M.; Server, A.C.; Nichols, R.A.; Shooter, E.M.  
Biochemistry 26, 2082-2085, 1987  
A:Title: Epidermal growth factor binding protein: identification of a different protein.  
A:Reference number: A27120; MUID:87299636; PMID:3304419  
A:Accession: A27120  
A:Molecule type: protein  
A:Residues: 25-54;112-124,'X',126-130;165-184,'X',186-187,'X',189-192 <ISA>  
R:Evans, B.A.; Drinkwater, C.C.; Richards, R.I.  
J. Biol. Chem. 262, 8027-8034, 1987  
A:Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of th  
A:Reference number: I55260; MUID:87250386; PMID:3036794  
A:Accession: I70015  
A:Status: preliminary; translated from GB/EMBL/DBDB  
A:Molecule type: DNA  
A:Residues: 70-122 <RES>  
A:Cross-references: GB:M18608; NID:g198500; PIDN:AAA39351.1; PID:g198506  
C:Comment: This sequence is one of approximately twenty-five members of a gene family on  
C:Genetics:  
A:Gene: KAL  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase; zymogen  
P:1-18/Domain: signal sequence #status predicted <SIG>  
P:19-24/Domain: propeptide #status predicted <PRO>  
P:25-261/Product: tissue kallikrein mCK-9, submandibular #status experimental <MAT>  
F:25-233/Domain: trypsin homology <TRI>  
F:65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:		
Pred. No.:	5,92e-42	Length:
Score:	582.00	Matches:
Percent Similarity:	59.23%	Conservative:
Best Local Similarity:	44.23%	Mismatches:
Query Match:	32.70%	Indels:
DB:	2	Gaps:
		3

[illegible]

472	Qy	ACGAGCCAGATGACCTTCCCTCTGACCTCATGTGCGGTGGAGTCAAGCTCATCTCC	531
161	Db	 ThrProPheLysPheGlnAsnAlaLysAspLeuGlnCysValAsnLeuLysLeuLeuPro	180
532	Qy	CCCAGGACTGCACGAAGGTTTCAAGAGACTTACTGGAAAAATTCATCTGTGCGCTCGC	591
181	Db	:::      AsnGluAspCysGlyLysAlaHisIleGluLysValThrAspValMetLeuCysAlaGly	200
592	Qy	ATCCCCGACTCCAAAGAAAACGCTGCATGGTGACTCAGGGGACCGTTGGTGTGCAGA	651
201	Db	 GluThrAspGlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAsp	220
652	Qy	GGTACCTCGAAGTCTGGTGCTCTCGGGAATTTCCCTTCGGGCAACCCCAATCACCCA	711
221	Db	 GlyValLeuGlnGlyIleThrSerTrpGlyPheThrProCysGlyGluProLysLysPro	240
712	Qy	GGAGTCTACACTCAAGTGTCAAGTTCCACCAAGTGGATAAATGACACCCATGAAAAGCAT	771
241	Db	:::      GlyValThrLysLeuIleLysPheThrSerTrpIleLysAspThrMetAlaLysAsn	260

## RESULT 4

S01971  
cisse kallikrein (EC 3.4.21.35) precursor - mouse  
N;Alternate names: glandular kallikrein  
C;Species: Mus musculus (house mouse)  
C;Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 09-Jul-2004  
C;Accession: S01971; I70023  
R;Drinkwater, C.C.; Richards, R.I.  
Nucleic Acids Res. 16, 10918, 1988  
A;Title: Sequence of mGK-11, a mouse glandular kallikrein gene.  
A;Reference number: S01971; MUID:89083511; PMID:3205728  
A;Accession: S01971  
A;Molecule type: DNA  
A;Residues: 1-261 <DRI>  
A;Cross-references: UNIPROT:P15946; EMBL:X13215; NID:G53064; PIDN:CAA31604.3; I70023  
R;Evans, B.A.; Drinkwater, C.C.; Richards, R.I.  
J. Biol. Chem. 262, 8027-8034, 1987  
A;Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis.  
A;Reference number: I55260; MUID:87250386; PMID:3036794  
A;Accession: I70023  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 70-122 <RES>  
A;Cross-references: GB:M18610; NID:G198529; PIDN:AAA39353.1; PID:G198537  
C;Genetics:  
A;Introns: 16/1; 69/2; 165/1; 210/3  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-24/Domain: propeptide #status predicted <PRO>  
F;25-261/Product: tissue kallikrein #status predicted <MAT>  
F;65-253/Domain: trypsin homology <TRY>  
F;65-120,213/Active site: His. Asp. Ser #status predicted

Alignment Scores:		
Pred. No.:	9,478-41	261
Score:	568.00	112
Percent Similarity:	57.9%	37
Best Local Similarity:	43.5%	928
Query Match:	31.91%	16
DB:	2	3
		Length:
		Matches:
		Conservative:
		Mismatches:
		Indels:
		Gaps:

US-09-905-083A-30 (1-969) x S01971 (1-261)

Qy	49	TTACTGCTATCCTTAGCCTTGGAAATCTGACGAGAGAGAGCCACG-----GGTAC	99
Db	4	LeuLeuPheLeuAlaLeuSerLeuGlyGlyLeuAlaAlaProProValGlnSer	23
Qy	100	AAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCACCATGCAGGTGCCTGCTC	159
Db	24	ArgLeuValGlyGlyPheAsnCysGluLysAsnSerGlnProTrpHisValAlaValTr	43











A:Molecule type: protein  
A:Residues: 165-174 <MUR>

C:Genetics:

A:Gene: mGK-6

A:Introns: 16/1; 69/2; 165/1; 210/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-261/Product: tissue kallikrein #status predicted <MAT>

F:25-253/Domain: tissue kallikrein #status predicted <MAT>

F:65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	2,74e-39	Length:	261
Score:	551.00	Matches:	105
Percent Similarity:	58.46%	Conservative:	47
Best Local Similarity:	40.3%	Mismatches:	92
Query Match:	30.9%	Indels:	16
DB:	2	Gaps:	3

US-09-905-083A-30 (1-969) x A25606 (1-261)

QY	40	CTGCAGATCTTACTCTCTTACCTTGGAACTGCAGGAGAGAGCCAG-----	93
DB	1	MetArgPheLeuLeuPheLeuAlaLeuSerLeuGlyGlyLeuAlaProPro	20
QY	94	---GGTGACAAGATTATTGATCGCGCCCATGTGCAAGAGGCTCCACCCATGCGCAGGTG	150
DB	21	ValGlnSerArgIleValGlyGlyPheAsnCysGluLysAsnSerGlnProTrpGlnVal	40
QY	151	GCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGCGCTCTGGTCAATGAGCGCTGGGTG	210
DB	41	AlaValTyArgPheThrTyrGlnCysGlyGlyLeuLeuAsnValAsnTrpVal	60
QY	211	CTCACTCGCGCCACTGCAAGATGAATGAGTACACCGTGCACCTCGGCGAGTATACGCTG	270
DB	61	LeuThrAlaAlaHisCysHisAsnAspTyrGlnValTrpLeuGlyLysAsnAsnPhe	80
QY	271	-----GGCGACGAGAGCTCAGAGGATCAAGGCTCGAAGTCAATTCGCCACCCCGGC	324
DB	81	LeuGluAspGluProSerAlaGlnHisArgLeuValSerLysAlaIleProHisProAsp	100
QY	325	TACTCC-----ACACGACCCATGTTAATGAC	351
DB	101	PheAsnMetSerLeuLeuAsnGluHisThrProGlnProGluAspAspTyrSerAsnAsp	120
QY	352	CTCATGCTCGTGAAGCTCAATAGCAGCGCAGGCTGCATCCATGGTGAAGAACTCAGG	411
DB	121	LeuMetLeuLeuArgLeuLysLysProAlaAspIleThrAspValValLysProIleAsp	140
QY	412	CTGCGCTCCCGCTGCGAACCCTCGAACCACTGTACTGTCTCGGCTGGGCGCACTACC	471
DB	141	LeuProThrGluGluProLysLeuGlySerThrCysLeuAlaSerGlyTrpGlySerIle	160
QY	472	ACGAGCCGATGTGACCTTTCCCTCTGACCTCATGTGGTGGATGTCAAGTCAATCTCC	531
DB	161	ThrProValTyThrGluTyProAspGluLeuGlnCysValAsnLeuLysLeuLeuPro	180
QY	532	CCCGAGGACTGCACCAAGGTTTACAAGACTTACTCGAAATTCATCTGTGTGGCTGGC	591
DB	181	AsnGluAspCysAlaLysAlaHisIleGluLysValThrAspAspMetLeuCysAlaGly	200
QY	592	ATCCCGGACTCCAAGAAAACCCCTGCAATGTGTGACTCAGGGGACCGTGTGTGTGCAGA	651
DB	201	AspMetAspGlyGlyAspThrCysAlaGlyAspSerGlyGlyProLeuIleCysAsp	220
QY	652	GGTACCTGCAAGGTCGTGTGTCGGGAATTTCCCTTGGCGGCAACCCCAATGACCA	711
DB	221	GlyValLeuGlnGlyIleThrSerTrpGlyProSerProCysGlyLysProAsnValPro	240
QY	712	GGAGTCTACACTCAAGTGTGCAAGTTCCCAAGTGGATAAATGACCATGAAAGCAT	771
DB	241	GlyIleTyThrArgValLeuAsnPheAsnThrTrpIleArgGluThrMetAlaGluAsn	260

RESULT 11

A37938

tissue kallikrein (EC 3.4.21.35), prostatic, precursor - dog

N:Alternate names: kallikrein homolog

C:Species: Canis lupus familiaris (dog)

C>Date: 21-Feb-1992 #sequence revision 21-Feb-1992 #text\_change 09-Jul-2004

C:Accession: A37938; S00613; A61555

R:Chapdelaine, P.; Gauthier, E.; Ho-Kim, M.A.; Bissonnette, L.; Tremblay, R.R.; Dube, J.

DNA Cell Biol. 10, 49-59, 1991

A:Title: Characterization and expression of the prostatic arginine esterase gene, a can

A:Reference number: A37938; MUID:91119675; PMID:1991049

A:Accession: A37938

A:Molecule type: DNA

A:Residues: 1-260 <CHA>

A:Cross-references: UNIPROT:P09582; GB:M63669; NID:q163906; PIDN:AAA30831.1; PID:g163907

R:Chapdelaine, P.; Ho-Kim, M.A.; Tremblay, R.R.; Dube, J.Y.

FEBS Lett. 232, 187-192, 1988

A:Title: Nucleotide sequence of the androgen-dependent arginine esterase mRNA of canine

A:Reference number: S00613; MUID:88211858; PMID:2835268

A:Accession: S00613

A:Molecule type: mRNA

A:Residues: 1-55, 'N' 57-260 <CH2>

A:Cross-references: EMBL:Y00751; NID:9867; PIDN:CAA68720.1; PID:g868

R:Chapdelaine, P.; Potvin, C.; Ho-Kim, M.A.; Larouche, L.; Bellemare, G.; Tremblay, R.T.

Mol. Cell. Endocrinol. 56, 63-70, 1988

A:Title: Androgen regulation of canine prostatic arginine esterase mRNA using cloned cDN

A:Reference number: A61555; MUID:88225749; PMID:3371547

A:Accession: A61555

A:Molecule type: mRNA

A:Residues: 105-260 <CH3>

A>Note: authors translated the codon GAG for residue 67 as Gln, and TTG for residue 97 a

C:Comment: This glandular kallikrein is not known to have a physiological substrate but

C:Superfamily: trypsin; trypsin homology

C:Keywords: extracellular protein; glycoprotein; hydrolase; serine proteinase

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-24/Domain: propeptide #status predicted <PRO>

F:25-252/Domain: trypsin homology <TRY>

F:25-97/Domain: tissue kallikrein, prostatic, light chain #status predicted <LCH>

F:25-97,98-260/Product: tissue kallikrein, prostatic #status predicted <MAT>

F:98-260/Domain: tissue kallikrein, prostatic, heavy chain #status predicted <HCH>

F:65,119,212/Active site: His, Asp, Ser #status predicted

F:779/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	5,49e-39	Length:	260
Score:	547.50	Matches:	108
Percent Similarity:	56.75%	Conservative:	35
Best Local Similarity:	42.86%	Mismatches:	96
Query Match:	30.76%	Indels:	13
DB:	2	Gaps:	3

US-09-905-083A-30 (1-969) x A37938 (1-260)

QY	49	TTACTGTCTATCTTAGCTTGGAACTCCAGAGAGAGAGCCAG---GGTGACAAGATT	105
DB	6	LeuCysLeuAlaMetSerLeuGlyTrpThrGlyAlaGluProHisPheGlnProArgile	25
QY	106	ATTGATGCGGCCCTTGTCGCAAGAGGCTCCACCCATGCGAGGTGGCTGCTCAGTGGC	165
DB	26	IleGlyGlyArgGluCysLeuLysAsnSerGlnProTrpGlnValAlaValTyHisAsn	45
QY	166	AATCAGCTCCACTGCGGAGGGCTCTGCTCAATGAGCGCTGGTGTCTCACTGCCGCCAC	225
DB	46	GlyGluPheAlaCysGlyGlyValLeuValHisProGluTrpValLeuThrAlaAlaHis	65
QY	226	TGCAAGATGAATGAGTACACCGTGCACCTGGCGAGTGTATACCTGGGCGACAGG---	279
DB	66	CysAlaAsnSerAsnCysGluValTrpLeuGlyArgHisAsnLeuSerGluSerGluAsp	85
QY	280	AGAGCTCAGAGGATCAAGGCGCTCGAAGTCAATTCCGCCACCCCGGCTACTCCACA---	333
DB	86	GluGlyGlnLeuValGlnValArgLysSerPheIleHisProLeuTyThrLysVal	105





Qy	97	GACAAGATTATGTGGCGCCCATGTGCAGAGAGGCTCCACCAATGGCAGGTGGCCCTG	156
Db	5	AspIleValGlyGlyTyrThrCysGlyAlaAsnThrValProTyrGlnValSerLeu	24
Qy	157	CTCAGTGGCAATCAGCTCCACTCCGAGGCGCTCGGTCAATGAGCGGTGGTGCTCACT	216
Db	25	AsnSerGlyTyrHisPhe---CysGlyGlySerLeuIleAsnSerGlnTrpValValSer	43
Qy	217	GCGGCCCATCTGCAAGATGAATGAGTACACCGTGACCTGGGGCAGTGATACGTGGGCG	273
Db	44	AlaAlaHisCysTyrLysSerGlyIleGlnValArgLeuGlyGluAspAsnIleAsnVal	63
Qy	274	--GACAGGAGAGCTCAGAGGATCAAGGCTCGAAGTCAATCCGGCACCCCGCTACTCC	330
Db	64	ValGluGlyAsnGluGlnPheIleSerAlaSerLysSerIleValHisProSerTyrAsn	83
Qy	331	ACACAGACCCATGTTAATGACTCATGTCGTGAAGCTCAATAGCACGCCACGGGTGCA	390
Db	84	SerAsnThrLeuAsnAsnAspIleMetLeuIleLysLeuLysSerAlaAlaSerLeuAsn	103
Qy	391	TCCATGTTGAAGAAAGTCAGGCTGCCCTCCCGCTCGGAACCCCTCGGACCACTGTACT	450
Db	104	SerArgValAlaSerIleSerLeuProThrSerCysAlaSerAlaGlyThrGlnCysLeu	123
Qy	451	GTCTCCGGCTGGGCACCTACACAGACCCAGATGTCACCTTTCTCCCTCTGACCTCATGTGC	510
Db	124	IleSerGlyTrpGlyAsnThrLysSerSerGlyThrSerTyrProAspValLeuLysCys	143
Qy	511	GTGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGTTTACAGGACTTACTGGAA	570
Db	144	LeuLysAlaProIleLeuSerAspSerSerCysLysSerAlaTyrProGlyGlnIleThr	163
Qy	571	AATTCATGCTGTGGCTGGCATCCCGACTCCAAGAAAACGCTGCAATGCTGACTCA	630
Db	164	SerAsnMetPheCysAlaGlyTyrLeuGluGlyGlyLysAspSerCysGlnGlyAspSer	183
Qy	631	GGGGACCGTTGGTGTGCAGAGGTACCTGCAAGGCTCTGGTGTCTCTGGGGAACCTTCCT	690
Db	184	GlyGlyProValValCysSerGlyLysLeuGlnGlyIleValSerTrpGlySer---Gly	202
Qy	691	TGCGGCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCCACCAAGTGGATA	750
Db	203	CysAlaGlnLysAsnLysProGlyValTyrThrLysValCysAsnTyrValSerTrpIle	222
Qy	751	AATGACACCATG	762
Db	223	LysGlnThrIle	226

RESULT 15

KQRTP

N;Alternate names: glandular kallikrein; kininogenin; true tissue kallikrein

C;Species: Rattus norvegicus (Norway rat)

C;Date: 05-Apr-1983 #sequence revision 05-Apr-1983 #text change 09-Jul-2004

C;Accession: A00944; A41429; A25137; JX0073; A23863; A33359

R;Swift, G.H.; Dagorn, J.C.; Ashley, P.L.; Cummings, S.W.; MacDonald, R.J.

Proc. Natl. Acad. Sci. U.S.A. 79, 7263-7267, 1982

A;Title: Rat pancreatic kallikrein mRNA: nucleotide sequence and amino acid sequence of

A;Reference number: A00944; MUID:83117659; PMID:6961406

A;Accession: A00944

A;Molecule type: mRNA

A;Residues: 1-265 <SWI>

A;Cross-references: UNIPROT:P00758

A;Experimental source: pancreatic

R;Kato, H.; Nakanishi, E.; Enryoji, K.; Hayaishi, I.; Oh-Ishi, S.; Iwanaga, S.

J. Biochem. 102, 1389-1404, 1987

A;Title: Characterization of serine proteinases isolated from rat submaxillary gland: w

A;Reference number: A41429; MUID:88198057; PMID:3482210

A;Accession: A41429

A;Status: preliminary

A;Molecule type: protein

A;Residues: 29-53, 'X', 55-87 <KAT>

R;Gerald, W.L.; Chao, J.; Chao, L.

Biochim. Biophys. Acta 866, 1-14, 1986  
A>Title: Immunological identification of rat tissue kallikrein cDNA and characterization  
A/Reference number: A25137; MUID:86131678; PMID:3004582  
A/Accession: A25137  
A/Molecule type: mRNA  
A/Residues: 115-265 <GER>  
R/Inoue, H.; Fukui, K.; Miyake, Y.  
J. Biochem. 105, 834-840, 1989  
A>Title: Identification and structure of the rat true tissue kallikrein gene expressed i  
A/Reference number: JX0073; MUID:89327211; PMID:2753879  
A/Accession: JX0073  
A/Molecule type: DNA  
A/Residues: 1-265 <INO>  
A/Cross-references: GB:D00448; NID:g220792; PIDN:BAA00346.1; PID:g220794  
A/Experimental source: kidney  
R/Ashley, P.L.; MacDonald, R.J.  
Biochemistry 24, 4512-4520, 1985  
A>Title: Kallikrein-related mRNAs of the rat submaxillary gland: nucleotide sequences of  
A/Reference number: A23863; MUID:86051477; PMID:2998455  
A/Accession: A23863  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-265 <ASH>  
A/Cross-references: GB:M11563; NID:g205029; PIDN:AAA41464.1; PID:g205030  
A/Experimental source: submaxillary gland  
R/Wines, D.R.; Brady, J.M.; Pritchett, D.B.; Roberts, J.L.; MacDonald, R.J.  
J. Biol. Chem. 264, 7653-7662, 1989  
A>Title: Organization and expression of the rat kallikrein gene family.  
A/Reference number: A33359; MUID:89214217; PMID:2708383  
A/Accession: A33359  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 5-265 <WLN>  
A/Cross-references: GB:M23874; GB:J04701; GB:M23875; GB:M23876; NID:g205007; PIDN:AAA414  
C/Comment: The kallikreins liberate lysyl-bradykinin, a vasoactive decapeptide, from kin  
C/Comment: The protein presumably assumes the two-chain form by cleavage between residu  
C/Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release l  
C/Genetics:  
A/Introns: 20/1; 73/2; 169/1; 214/3  
C/Superfamily: trypsin; trypsin, trypsin, trypsin  
C/Keywords: hydrolase; pancreas; serine proteinase; zymogen  
F/1-17/Domain: signal sequence #status predicted <SIG>  
F/18-28/Domain: activation peptide #status predicted <APT>  
F/29-265/Product: tissue kallikrein, pancreatic #status predicted <MPT>  
F/29-257/Domain: trypsin homology <TRY>  
F/35-177, 54-70, 156-223, 198-202, 213-238/Disulfide bonds: #status predicted  
F/69, 124, 217/Active site: His, Arg, Ser #status predicted

```
QY 280 AGA-----GCTCAGAGGATCAAGCCTCGAAGTCATTCGCCACCCCGGCTACTCCACA 333
Db      |||||
88 GluProPheAlaGlnHisArgLeuValSerGlnSerPheProHisProGlyPheAsnGln 107
      |||||
QY 334 -----CAGACCCATGTTAATGACCTCATGCTC 360
Db      :: |||||
108 AspLeuIleTrpAsnHisThrArgGlnProGlyAspAspTyrSerAsnAspLeuMetLeu 127
      :: |||||
QY 361 GTGAAGCTCAATAGCCAGCCAGGCTGTCATCCATGGTGAAGAAAGTACAGGCTGCCCTCC 420
Db      |||||
128 LeuHisLeuSerGlnProAlaAspIleThrAspGlyValIleAspLeuProIle 147
      |||||
QY 421 CGCTGCGAAACCCCTGGAAACCACTGTCTCCGGCTGGGGCACTACCAAGAGCCCA 480
Db      |||||
148 GluGluProLysValGlySerThrCysLeuAlaSerGlyTrpGlySerIleThrProAsp 167
      |||||
QY 481 GATGTGACCTTTCCCTCTGACCTATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGAC 540
Db      |||||
168 GlyLeuGluLeuSerAspAspLeuGlnCysValAsnIleAspLeuLeuSerAsnGluLys 187
      |||||
QY 541 TGCACGAAGTTTACAAGGACTTACTGGAAATTCATGCTGTGGCTGGCATCCCCGAC 600
Db      |||||
188 CysValGluAlaHisLysGluGluValThrAspLeuMetLeuCysAlaGlyGluMetAsp 207
      |||||
QY 601 TCCAAAGAAAACGCTGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCCTG 660
Db      |||||
208 GlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAsnGlyValLeu 227
      |||||
QY 661 CAAGTCTGGTCTCTGGGGAACCTTCCCTTGGGCGCAACCCCAATGACCCAGAGTCTTAC 720
Db      |||||
228 GlnGlyIleThrSerTrpGlyPheAsnProCysGlyGluProLysLysProGlyIleTyr 247
      |||||
QY 721 ACTCAAGTCTGCAAGTTACCAAGTGGATAAATGACCATGAAAAAGCAT 771
Db      |||||
248 ThrLysLeuIleLysPheThrProTrpIleLysGluValMetLysGluAsn 264
      |||||
```

Search completed: July 14, 2005, 17:33:33  
Job time : 50.5 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 14, 2005, 17:07:08 ; Search time 203.5 Seconds  
(without alignments)  
4876.705 Million cell updates/sec

Title: US-09-905-083A-30  
Perfect score: 1780  
Sequence: 1 ggatttcgggctccatggc.....aagaacacacacccctcag 969

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09905083/runat\_14072005\_140352\_10610/app\_query.fasta\_1.1159  
-DB=Uniprot -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09905083@cgn\_1\_244@runat\_14072005\_140352\_10610 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1364	76.6	253	1 KLK7_HUMAN	P49862 homo sapien
2	1353	76.0	253	2 Q8NSN9	Q8NSN9 homo sapien
3	1028.5	57.8	249	2 Q91VE3	Q91VE3 m thymopain
4	988	55.5	181	2 Q8NFV7	Q8NFV7 homo sapien
5	951.5	53.5	234	2 Q9F048	Q9F048 mus musculus
6	614	34.5	260	1 NRPN_MOUSE	Q61955 mus musculus
7	612.5	34.4	293	1 KLK5_HUMAN	Q9Y337 homo sapien
8	611.5	34.4	260	1 NRPN_RAT	Q88780 rattus norv
9	609.5	34.2	276	2 Q8CGR6	Q8CGR6 mus musculus
10	604.5	34.0	246	2 Q6P3Z0	Q6P3Z0 mus musculus
11	593.5	33.3	250	2 Q8CGR5	Q8CGR5 mus musculus
12	589	33.1	293	2 Q9D140	Q9D140 mus musculus
13	586.5	32.9	242	2 Q8QV54	Q8QV54 mus musculus
14	583	32.8	251	1 KLKE_HUMAN	Q9P0G3 homo sapien
15	583	32.8	251	2 Q6B089	Q6B089 homo sapien
16	582.5	32.7	260	1 KLK8_HUMAN	Q60259 homo sapien

17	582	32.7	261	1 KLK9_MOUSE	P15949 mus musculus
18	580.5	32.6	260	2 Q8IW69	Q8IW69 homo sapien
19	578	32.5	277	1 KLIK_HUMAN	Q9UKR3 homo sapien
20	576.5	32.4	239	2 Q83275	Q83275 rattus norv
21	573.5	32.2	255	2 Q7JIG6	Q7JIG6 saguinus oe
22	571.5	32.1	250	1 KLIK_MOUSE	Q9UBX7 homo sapien
23	568	31.9	261	1 KLIK_MOUSE	P15946 mus musculus
24	567.5	31.9	254	1 KLK4_HUMAN	Q9Y5K2 homo sapien
25	564.5	31.7	250	2 Q632F2	Q632F2 rattus norv
26	563	31.6	256	1 KLF_MOUSE	Q9ZIR5 homo sapien
27	563	31.6	261	1 KLIK_MOUSE	P00756 mus musculus
28	563	31.6	261	1 KLIK_RAT	P63774 rattus norv
29	561.5	31.5	255	2 Q6ISIO	Q6ISIO homo sapien
30	561.5	31.5	276	2 Q9QYN3	Q9QYN3 m hippostas
31	561	31.5	246	1 TRY1_RAT	P0762 rattus norv
32	560	31.5	263	1 KLIK_MOUSE	Q9JN71 mus musculus
33	559.5	31.4	255	2 Q96RQ0	Q96RQ0 homo sapien
34	556	31.2	275	2 Q8IXD7	Q8IXD7 homo sapien
35	555	31.2	248	1 KLIK_HUMAN	Q9UKR0 homo sapien
36	554.5	31.2	234	2 Q9CV76	Q9CV76 mus musculus
37	553.5	31.1	249	2 Q9QYN4	Q9QYN4 mus musculus
38	552.5	31.0	260	1 ESTA_CANPA	P09582 canis famill
39	552	31.0	261	1 KLIK_MOUSE	P00755 mus musculus
40	552	31.0	261	1 KLIK_MOUSE	P15947 mus musculus
41	552	31.0	261	1 KLIK_RAT	P36373 rattus norv
42	551.5	31.0	246	2 Q6IE66	Q6IE66 rattus norv
43	551	31.0	261	2 Q8C232	Q8C232 mus musculus
44	551	31.0	276	1 KLIK_HUMAN	Q43240 homo sapien
45	548	30.8	263	1 KLIK_MOUSE	Q61754 mus musculus

ALIGNMENTS

RESULT 1  
ID KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RC TISSUE=Skin;  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
EL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Keratinocytes;  
RA Yousef G.M., Scorillas A., Diamandis E.P.;  
RT "Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene";  
RL Submitted (DSC-1999) to the EMBL/GenBank/DBSJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepers B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.

RA Hanson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrant P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
 RT mice; a model for chronic itchy dermatitis.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]

# CHARACTERIZATION.

RX MEDLINE=95314630; PubMed=7794273;

RA Skyt A., Stroemqvist M., Egelrud T.;

RT "Primary substrate specificity of recombinant human stratum corneum

RT chymotryptic enzyme."; Commun. 211:586-589(1995).

RL Biochem. Biophys. Res. Commun. 211:586-589(1995).

CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive

CC structures in the cornified layer of the skin in the continuous

CC shedding of cells from the skin surface. Specific for amino acid

CC residues with aromatic side chains in the P1 position. SCCE

CC cleaves insulin B chain at 6-Leu-Cys-7, 16-Tyr-Leu-17, 25-

CC Phe-Tyr-26, and 28-Tyr-Thr-27. Could play a role in the

CC activation of precursors to inflammatory cytokines.

CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is

CC expressed by keratinocytes in the epidermis. Very low levels are

CC also seen in the brain and kidney.

CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein

CC subfamily.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----

DR EMBL; L33404; AAC37551.1; -;

DR EMBL; AF166330; RAD49718.1; -;

DR EMBL; AF243527; AAK33360.1; -;

DR EMBL; AF332583; AAK69624.1; -;

DR PIR; A53968; A53968.

DR HSSP; P00760; 1E2K.

DR MEROPS; S01.300; -;

DR Genew; HGNC:6368; KLK7.

DR H-invDB; HIX0015373; -;

DR MIM; 604438; -;

DR GO; GO:0008236; F:serine-type peptidase activity; TAS.

DR GO; GO:0008544; P:epidermal differentiation; TAS.

DR InterPro; IPR009003; Pept\_Ser\_Cys.

DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1a.

DR Pfam; PF00089; Trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS00240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

DR Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;

KW Signal; Zymogen.

FT SIGNAL 1 22

FT PROPEP 23 29 Activation peptide.

FT CHAIN 30 253 Kallikrein 7.

FT ACT\_SITE 70 70 Charge relay system (By similarity).

FT ACT\_SITE 112 112 Charge relay system (By similarity).

FT ACT\_SITE 205 205 Charge relay system (By similarity).

FT DISULFID 36 137 By similarity.

FT DISULFID 55 71 By similarity.

FT DISULFID 71 239 By similarity.

FT DISULFID 144 211 By similarity.

FT DISULFID 176 190 By similarity.

FT DISULFID 201 226 By similarity.

FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).

SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;

Alignment Scores:

Pred. No.: 6.12e-110 Length: 253

Score: 1364.00 Matches: 253  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 76.63% Indels: 0  
 DB: 1 Gaps: 0

US-09-905-083A-30 (1-969) x KLK7\_HUMAN (1-253)

QY 16 ATGGCAACATCCCTTCTCTCCCTGCAGATCTTACTGTCTATCTTAGCCTTGAAACT 75  
 DB 1 MetAlaArgSerLeuLeuLeuProLeuGlnLeuLeuLeuSerLeuAlaLeuGluThr 20

QY 76 CAGAGAGAAGAGCCAGGGTGACAGATTATTGATGGCCCCCATGTCGCAAGAGCTCC 135  
 DB 21 AlaGlyGluGluAlaGlnGlyAspLysIleileAspGlyAlaProCysAlaArgGlySer 40

QY 136 CACCCATGGCAGGTGGCCCTCTCAGTGGCAATCAGCTCCACTGGGAGGGCTCTGCTC 195  
 DB 41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

QY 196 AATGAGCGCTGGGTGCTCACTGCCGCCACATGCAAGATGAATGATCACCCTGCACCTG 255  
 DB 61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

QY 256 GCGAGTGATACGCTGGGCGACAGAGAGCTCAGAGGCTCAAGGGCTCGAAGTCATTCGCG 315  
 DB 81 GlySerAspThrLeuGlyAspArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

QY 316 CACCCCGGCTACTCCACAGACACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATGCG 375  
 DB 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

QY 376 CAGGCGAGCTGTCTATCCATGCTGAAGAAAGTCAAGGCTCCCTCCGCTGGGAACCCCT 435  
 DB 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

QY 436 GGAACCACTGTACTGTCTCCGGCTGGGGCACTACACAGAGCCAGATGTGACCTTTCC 495  
 DB 141 GlyThrThrCysThrValSerGlyTyrGlyThrThrThrSerProAspValThrPhePro 160

QY 496 TCTGACCTCATGTGGTGGATGTCAGCTCATCTCCCTCCCGAGGACTGCAAGAGTTTAC 555  
 DB 161 SerAspLeuMetCysValAspValLysLeuLysSerProGlnAspCysThrLysValTyr 180

QY 556 AAGACTTACTGGAAATTTCCATGCTGTGCTGGCATCCCGACTCCCAAGAAAAACGCC 615  
 DB 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAla 200

QY 616 TGCATGTGTACTCAGGGGGACCGTTGTGTGTCAGAGGTACCTGCAAGGTCTGTGTGCC 675  
 DB 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

QY 676 TGGGGACTTTCCCTTGGCGCCCAACCCCAATCACCAGGAGTCTACACCAAGTGTGCAAG 735  
 DB 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

QY 736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774

DB 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

## RESULT 2

Q8N5N9

ID Q8N5N9 PRELIMINARY; PRT; 253 AA.

AC Q8N5N9;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Stratum corneum chymotryptic enzyme, preproprotein.

GN Name=KLK7;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;



RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor. C3;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvJ;  
RA Hanson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
RA Wallbrant P., Egelrud T.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA The FANTOM Consortium;  
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [9]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [10]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akai S., Tanaka T., Tanaka T.,  
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA Adachi J., Aizawa K., Akimura T., Aono H., Arai A.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuura T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tanaka A., Takahashi F., Tanaka T.,  
RA Teima Y., Toyota T., Yamamura T., Yamanaka I., Yasunishi A.,  
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AB008371; BAB55604.1; -  
DR EMBL; BC027823; AAH27823.1; -  
DR EMBL; AF339930; AAK69652.1; -  
DR EMBL; AK029477; BAC26467.1; -  
DR EMBL; AK077406; BAC36787.1; -  
DR MEROPS; P00760; 1EZX.  
DR MGD; MGI:1346336; K1K7.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR Pfam; PF00089; Trypsin; 1  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 249 AA; 27257 MW; 0D4E380F12D14F87 CRC64;  
  
Alignment Scores:  
Pred. No.: 9,82e-81 Length: 249  
Score: 1028.50 Matches: 185  
Percent Similarity: 85.48% Conservative: 27  
Best Local Similarity: 74.60% Mismatches: 33  
Query Match: 57.78% Indels: 3  
DB: 2 Gaps: 1  
US-09-905-083A-30 (1-969) x Q91VE3 (1-249)  
  
QY 31 CTCCTGCCCCGTGACAGATCTTACTGCTATCTCTTACCTTGGAACTGCGAGGAGAGAGGCC 90  
Db 5 LeuLeuSerLeuLeuThrValLeuLeuSerLeuAlaLeuGluThrAlaGly----- 21  
  
QY 91 CAGGTCACACAGATTATTGATGGCGCCCATGTCACAGAGGCTCCACCCATGCGAGCTG 150  
Db 22 GlnGlyGluArgIleIleAspGlyTyrIysCysIysGluGlySerHisProTyrGlnVal 41  
  
QY 151 GCCCTGCTCAGTGCACATCAGCTCCACTGCGGAGCGCTCTGGTCAATAGCGCTGGGTG 210  
Db 42 AlaLeuLeuIysGlyAsnGlnLeuHisCysGlyGlyValLeuValAspLysTyrTrpVal 61  
  
QY 211 CTGACTGCGCCCACTGCAAGATCAATGAGTACACCTGCGGAGGAGTACGCTG 270  
Db 62 LeuThrAlaAlaHisCysLysMetGlyIleThrGlnValGlnLeuGlySerAspLysIle 81  
  
QY 271 GGCCACAGGAGAGCTCAGAGGATCAAGGCTTCGAGGATCATTCGCCGCCCGGCTACTCC 330  
Db 82 GlyAspGlnSerAlaGlnLysIleLeuAlaThrLysSerPheArgHisProGlyTyrSer 101



DR GO:0005615; C:extracellular space; TAS.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 234 234  
 SQ SEQUENCE 234 AA; 25220 MW; BEBE57D6C325B85F CRC64;

Alignment Scores:  
 Pred. No.: 4, 896-74 Length: 234  
 Score: 951.50 Matches: 175  
 Percent Similarity: 84.98% Conserved: 23  
 Best Local Similarity: 75.11% Mismatches: 32  
 Query Match: 53.46% Indels: 3  
 Dbs: 2 Gaps: 1

US-09-905-083a-30 (1-969) x Q9R048 (1-234)

QY 31 CTCCTGCCCTCCAGATCTTACTGCTATCTTGGAACTGCGAGAGAGGCC 90  
 Db 5 LeuLeuSerLeuThrValLeuLeuSerLeuAlaLeuGluThrAlaGly----- 21  
 QY 91 CAGGTGACAGATTATTGATGGCCGCCCATGTGCAAGAGGCTCCACCCATGGCAGTG 150  
 Db 22 GlnGlyGluArgIleAspGlyTyrLysCysLysGluGlySerHisProTyrpGlnVal 41  
 QY 151 GCCTGCTCAGTGGCAATCAGCTCCACATCGCGAGCGCTCGTGTCAATGAGCGCTGGTG 210  
 Db 42 AlaLeuLeuLysGlyAsnGlnLeuHisCysGlyGlyValLeuValAspLysTyrTrpVal 61  
 QY 211 CTCCTGCGCCCATCGCAAGATGAATGATGACCTGCACCTGGGCGAGTGATACGTG 270  
 Db 62 LeuThrAlaAlaHisCysLysMetGlyGlnTyrGlnValGlnLeuGlySerAspLysIle 81  
 QY 271 GCGCACAGAGNCTCAGAGNATCAAGCTCGAAGTATTCCTCCGCCACCCCGCTACTCC 330  
 Db 82 GlyAspGlnSerAlaGlnLysIleLysAlaThrLysSerPheArgHisProGlyTyrSer 101  
 QY 331 ACACAGACCCATGTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGCGAGCTGTCA 390  
 Db 102 ThrLysThrHisValAsnAspIleMetLeuValArgLeuAspGluProValLysMetSer 121  
 QY 391 TCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTCGAACCCTCGGAACCACTGTACT 450  
 Db 122 SerLysValGluAlaValGlnLeuProGluHisCysGluProProGlyThrSerCysThr 141  
 QY 451 GTCTCCGCTGGGCACTACACGAGCCAGATGTGACCTTTCCTCTGACCTCATGTGC 510  
 Db 142 ValSerGlyTyrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMetCys 161  
 QY 511 GTGGATGTCAGCTCATCTCCCGCCAGACTGCACGAAGGTTTCAAGCACTTACTGAA 570  
 Db 162 SerAspValLysLeuLysSerSerArgLysLysValLysValLysAspLeuLeuGly 181  
 QY 571 AATTCATGCTGTGCGTGGCATCCCGACTCCAAAGAAAACGCTCGAATGGTGACTCA 630  
 Db 182 LysThrMetLeuCysAlaGlyIleProAspSerLysThrAsnThrCysAsnGlyAspSer 201  
 QY 631 GGGGACCGTGGTGTGACAGGTACCTGACAGTCTGGTGTCTCTGGGGAACCTTCCT 690  
 Db 202 GlyGlyProLeuValCysAsnAspThrLeuGlnGlyLeuAlaSerArgGlyThrTyrPro 221  
 QY 691 TGGCGCCCAACCAATGACCCAGAGCTTACACTCAAGTG 729  
 Db 222 CysGlyGlnProAsnAspProGlyValTyrThrGlnVal 234

RESULT 6

NRPN MOUSE  
 ID NRPN MOUSE STANDARD; PRT; 260 AA.  
 AC Q61955;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Neuropsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8).  
 GN Name=Klk8; Synonyms=Nrpn, Prss19;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Hippocampus;  
 RX MEDLINE=9534817; PubMed=7623137;  
 RA Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,  
 RA Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.;  
 RT "Expression and activity-dependent changes of a novel limbic-serine  
 RT protease gene in the hippocampus.";  
 RL J. Neurosci. 15:5088-5097(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Yoshida S., Hirata A., Inoue N., Shiosaka S.;  
 RT "Cloning and assignment of mouse neuropsin gene, Prss19 to chromosome  
 RT 7B4.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Czech II; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF N-TERMINUS, CHARACTERIZATION, AND MASS SPECTROMETRY.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=98225202; PubMed=9556608; DOI=10.1074/jbc.273.18.11189;  
 RA Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,  
 RA Shiosaka T., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;  
 RT "Characterization of recombinant and brain neuropsin, a plasticity-  
 RT related serine protease.";  
 RL J. Biol. Chem. 273:11189-11196(1998).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=95134351; PubMed=9933620; DOI=10.1074/jbc.274.7.4220;  
 RA Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,  
 RA Shiosaka S., Hakoshima T.;  
 RT "Crystal structure of neuropsin, a hippocampal protease involved in  
 RT kindling epileptogenesis.";  
 RL J. Biol. Chem. 274:4220-4224(1999).  
 CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
 CC hippocampal plasticity. Has a strong proteolytic activity against  
 CC fibronectin.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.

```

-|- ENZYME REGULATION: Strongly inhibited by diisopropyl
fluorophosphate, leupeptin and (4-aminophenyl)methanesulfonyl 1-
fluoride.
-|- SUBCELLULAR LOCATION: Secreted.
-|- TISSUE SPECIFICITY: Expressed specifically in the limbic system of
mouse brain and is localized at highest concentration in pyramidal
neurons of the hippocampal CA1-3 subfields.
-|- MASS SPECTROMETRY: MW=26613; METHOD=MALDI; RANGE=29-260;
NOTE=Ref. 4.
-|- MASS SPECTROMETRY: MW=26229; METHOD=MALDI; RANGE=33-260;
NOTE=Ref. 4.
-|- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

```

[illegible]









DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Jaw and Limb;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriques S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skaleja U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Jaw and Limb;  
RA Strausberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; BC063763; AAH63763.1; -;  
DR HSP; P00761; IAKS.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYPSIN\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase, Hypothetical protein; Protease; Serine protease.  
SQ SEQUENCE 246 AA; 27213 MW; 0BC1454D6CC8CDEC CRC64;  
Alignment Scores:  
Pred. No.: 7,98e-44 Length: 246  
Score: 604.50 Matches: 112  
Percent Similarity: 67.11% Conservative: 39  
Best Local Similarity: 49.78% Mismatches: 67  
Query Match: 33.96% Indels: 7  
DB: 2 Gaps: 4  
US-09-905-083A-30 (1-969) x Q6P3Z0 (1-246)  
QY 112 GCGCGCCCATGTGCAAGAGGCTCCACCCATGCGAGGTGGCCCTGCTCAGTGGCAATCAG 171  
Db 10 GlyThrCysLeuProHisSerGlnProTglnAlaLeuLeuIleargGlyarg 29  
QY 172 CTCCTACTGGAGGGCTCTGTCATATGAGCGCTGGTCTCACTGGCGCCCATGCAAG 231  
Db 30 LeuLeuCysGlyGlyValLeuValHisProLysTrpValLeuThraAlaHisCysarg 49  
QY 232 ATGAATGAGTACACCGTGCACCTCGGCGAGTGATACCTGGGC-----GACAGGAGAGCT 285

Db 50 LysaspGlyThrValHisLeuGlyLysHisAlaLeuGlyArgValGluasnGlyGlu 69  
QY 286 CAGAGGATCAAGGCTCGAAGTCAATTCGCGCACCCCGGCTACTCC---ACACAGACCCAT 342  
Db 70 GlnAlaMetGluValValArgSerIleProHisProGluThrGlnValThrProThrHis 89  
QY 343 GTTAAT-----GACCTCATGCTGCTGAAGCTCAATAGCCAGGCGCCAGCTGTCATCC 393  
Db 90 LeuasnHisAspHisAspIleMetLeuLeuGluLeuLysSerProValGlnLeuSerSer 109  
QY 394 ATGTGTGAAGAAGTCAAGCTGCTCCCTCC---CGCTGCGAAGACCCCTGGGAACACCTGTACT 450  
Db 110 HisValargThrLeuLysLeuSerAlaAspCysLeuProThrGlyThrCysCysarg 129  
QY 451 GTCTCCGCTGGGCGCATCACAGAGCCAGATGTGACCTTTCCCTCTGACCTCATGTGC 510  
Db 130 ValSerGlyTrpGlyThrThrSerProGlnValAsnTrpProLysThrLeuGlnCys 149  
QY 511 GTGGATGTCAAGCTCATCTCCCGCCAGGACTGCAGAGGTTTACAAGGACTTACTGGAA 570  
Db 150 AlaasnIleGluLeuArgSerAspGluGlnCysargGlnValTrpProGlyLysIleThr 169  
QY 571 AATTCATGCTGTGCTGGCATCCCGGACTCCCAAGAAAAACGCTGCAATGGTGACTCA 630  
Db 170 AlaasnMetLeuCysAlaGlyThrLysGluGlyLysAspSerCysGluGlyAspSer 189  
QY 631 GGGGACCGTGTGTGTCAGAGGTACCTGCAAGGTCTGTGTCTGGGGAACTTTTCCT 690  
Db 190 GlyGlyProLeuIleCysasnGlyLysLeuTrpGlyLysIleSerTrpGlyAspPhePro 209  
QY 691 TCGCGCCCAACCAATGACCCAGGAGTCTTACACTCAAGTGTGCAAGTTCACCAAGTGGATA 750  
Db 210 CysGlyGlnProAsnArgProGlyValTrpThrArgValSerLysTrpLeuArgTrpIle 229  
QY 751 AATGACACCATGAAA 765  
Db 230 ArgGluIleIleArg 234

RESULT 11  
Q8CGR5  
ID Q8CGR5 PRELIMINARY; PRT; 250 AA.  
AC Q8CGR5;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Glandular kallikrein KLK14.  
GN Name=KLK14;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22325484; PubMed=12437987; DOI=10.1016/S0006-291X(02)02629-3;  
RA Olsson A.Y., Lundwall A.;  
RT "Organization and evolution of the glandular kallikrein locus in Mus musculus.";  
RL Biochem. Biophys. Res. Commun. 299:305-311 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M., Mural R.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AY152433; AAN78421.1; -;  
DR HSP; P00760; IEZX.  
DR MEROPS; S01.029; -;  
DR MGD; MGI:2447564; Klk14.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_s1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; TRYPSIN; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 250 AA; 27016 MW; F62PE8F2290FBE8 CRC64;

Alignment Scores:  
Pred. No.: 7,27e-43 Length: 250  
Score: 593.50 Matches: 114  
Percent Similarity: 61.66% Conservative: 42  
Best Local Similarity: 45.06% Mismatches: 86  
Query Match: 33.34% Indels: 11  
DB: 2 Gaps: 4

US-09-905-083A-30 (1-969) x Q8CGR5 (1-250)

QY 28 CTTCTCCTCCCTCCAGATCTTACTGCTATCTTAGCTTGGAACTGCAGGAGA 87  
DB 4 LeuLeuIleLeuGlnAlaLeuAlaValAlaIleAla-----Gln 17  
QY 88 GCCCAGGGTGAC---AAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCACCCATGG 144  
DB 18 SerGlnGlyAspHisIleIleGlyGlyTyrArgCysValArgAsnSerGlnProTyr 37  
QY 145 CAGGTGGCCCTGCTCAGT-----GGCAATCAGCTCCACTGGGAGGCGTCTGTGTCAT 198  
DB 38 GlnValAlaLeuGlnAlaGlyProGlyHisArgPheLeuCysGlyGlyValLeuLeuSer 57  
QY 199 GAGCGCTGGTCTCAGTCCGCCCATCTGCAAGATGAATGATCAGCTGCCTGGCC 258  
DB 58 AspGlnTrpValIleThrAlaAlaHisCysAlaArgProIleLeuHisValAlaLeuGly 77  
QY 259 AGTGATACGCTG-----GGCCAGGAGAGCTCAGAGGATCAAGCGCTCGAAGTCATTC 312  
DB 78 LysHisAsnIleArgArgTrpGluAlaThrGlnGlnValArgValAlaArgGlnVal 97  
QY 313 CCCACCCCGGCTACTCCACAGACCACTGTAATGATCCTCATCTCTGTAAGTCAAT 372  
DB 98 ProHisProGlnTyrGlnProGlnAlaHisAspAsnAspLeuMetLeuLeuLysLeuGln 117  
QY 373 ACCCAGGCGAGCTGTCATCCATGCTGAAGAAGTCAGGCTGCCCTCCGCTGCGAACCC 432  
DB 118 LysLysValArgLeuGlyArgAlaValLysThrIleSerValAlaSerSerCysAlaSer 137  
QY 433 CCGTGAACCCACTGTACTGTCTCCGGCTGGGGCACTACCAGCGCCAGATGTGACCTTT 492  
DB 138 ProGlyThrProCysArgValSerGlyTrpGlyThrIleAlaSerProIleAlaArgTyr 157  
QY 493 CCTCTGACCTCATGTGGTGATATCAAGTCAATCTCCGCCAGGATCGACGAAGTT 552  
DB 158 ProThrAlaLeuGlnCysValAsnValAsnIleMetSerGlnGlnAlaCysHisArgAla 177  
QY 553 TACAGGACTTACTGGAAATTCATGCTGCTGGCTGGCATCCCGACTCCAGAAAC 612  
DB 178 TyrProGlyIleIleThrSerGlyMetValCysAlaGlyValProGluGlyGlyLysAsp 197  
QY 613 GCCTGCAATGTGTACTCAGGGGACCGTTGTGTGTGTCAGAGGTGTACCTGCAAGGTGTGTG 672  
DB 198 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlyGlyGlnGlnGlyLeuVal 217  
QY 673 TCCTGGGAACTTTCCTTGGGGCCACCAATGATCCAGAGGTGTACACTCAAGTGTGC 732  
DB 218 SerTrpGlyMetGluArgCysAlaMetProGlyTyrProGlyValTyrAlaAsnLeuCys 237  
QY 733 AAGTTTCAACCAAGTGTGATAATGACACCATGATGAAGCAT 771  
DB 238 AsnTyrHisSerTrpIleGlnArgThrMetGlnSerAsn 250

RESULT 12  
Q9D140  
ID Q9D140 PRELIMINARY; PRT; 293 AA.  
AC Q9D140;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched  
DE library, clone:1110030019 product:weakly similar to KALLIKREIN 5 (BC  
DE 3.4.21.-) (STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN 2)  
DE (KUK-L2).  
GN Name=1110030019rik;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Mech. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA The FANTOM Consortium;  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitzunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

Db      255 LysLeuGlnGlyLeuValSerTrpIleYAspPheProCysAlaGlnArgAsnArgProGly 275  
Qy      715 GTCTACACTCAAGTGTGCAAGTTCCACCAAGTGATGAATAATGCACCACATGAAAAAGCAT 771  
              :::|||||              |||              |||              |||  
Db      275 ValtyrThrAsnLeuCysGluPheValIysTrpIleYAspThrMetAsnSerAen 293  
  
RESULT 13  
QBVS4      PRELIMINARY;      PRT;      242 AA.  
ID      Q8OVSA4  
AC      Q8OVSA4  
DT      01-JUN-2003 (TrEMBLrel. 24, Created)  
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE      Klk14 protein (Fragment).  
GN      Name=Klk14;  
OS      Mus musculus (Mouse).  
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX      NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=NMR1; TISSUE=Mammary tumor;  
MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA      Strausberg R.L., Feingold B.A., Grouse L.H., DeRge J.G.,  
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA      Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA      Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA      Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA      Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA      Krzywinski M.I., Skalska U., Smalls D.E., Schermer A., Schein J.E.,  
RA      Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences."  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=NMR1; TISSUE=Mammary tumor;  
RA      Strausberg R.;  
RA      Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
CC      -1- SIMILARITY: Belongs to peptidase family S1.  
DR      EMBL; BC044756; AAA44756.1; --  
DR      HSSP; P00760; 1BZX.  
DR      MEROPS; S01\_029; --  
DR      MGd; MG12447564; Klk14.  
DR      GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR      GO; GO:0008233; F:peptidase activity; IEA.  
DR      GO; GO:0004295; F:trypsin activity; IEA.  
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR      InterPro; IPR001254; Peptidease\_S1.  
DR      InterPro; IPR001314; Peptidease\_SIA.  
DR      InterPro; IPR009003; Pept\_Ser\_Cys.  
DR      Pfam; PF00089; Trypsin; 1.  
DR      PRINTS; PR00722; CHYMOTRYPSIN.  
DR      SMART; SM00020; TRYD\_SPC; 1.  
DR      PROSITE; PS02040; TRYPsin\_DOM; 1.  
DR      PROSITE; PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
DR      PROSITE; PS00135; TRYPsin\_SER; 1.  
KW      Hydrolase; Protease; Serine protease.  
FT      NON TER 1  
SQ      SEQUENCE 242 AA; 26059 MW; AJF8AG24DE481D36 CRC64;

```
Percent Similarity: 61.83% Conservative: 38
Best Local Similarity: 46.06% Mismatches: 87
Query Match: 32.95% Indels: 5
DB: 2 Gaps: 3

US-09-905-083a-30 (1-969) x Q80VS4 (1-242)

QY 64 GCCTTGGAACTGCAGGAGAAAGCCCGAGGTGAC--AAAGATTATTGATGGCGCCCA 120
DB 2 AlaLeuAlaValAlaLeuAlaGlnSerGlnGlyAspHisGlySilelleGlyTyrArg 21
QY 121 TGTGCAAGAGGTCCACACCATGGAGGTGGCCCTGCTCAGT-----GGCATCAGCTC 174
DB 22 CysValArgAsnSerGlnProTrpGlnValAlaLeuGlnAlaGlyProGlyHisArgPhe 41
QY 175 CACTCGGAGGGCTCTGTCAATCAGCGCTGGGTGCTCACTGCCGCCCATCGCAAGATG 234
DB 42 LeucysGlyValLeuLeuSerAspGlnTrpValIleThrAlaAlaHisCysAlaArg 61
QY 235 AATGAGTACACCGTGCACCTGGCGAGTATGATCGTGTG-----GGCGACAGGAGAGCTCAG 288
DB 62 ProfileLeuHisValAlaLeuGlyLysHisAsnIleArgTgTgPAlaAlaThrGlnGln 81
QY 289 AGGATCAAGGCTCGAAGTCATTCGCCACCCCGCTACTCAGACACCATCGTTTAT 348
DB 82 ValValArgValAlaArgGlnValProHisProGlnTrpGlnProGlnAlaHisAspAsn 101
QY 349 GACCTCATGCTCGTGAAGTCAATAGCAGCGCCAGCTGTCATCCATGCTGAAGAAAGTC 408
DB 102 AspLeuMetLeuLeuLysValGlnLysValArgLeuGlyArgAlaValIsthThrIle 121
QY 409 AGGCTGCCCTCCCGTGCAGAACCCCTCGAACCACTGTACTGTCTCCGGCTGGGGCACT 468
DB 122 SerValAlaSerSerCysAlaSerProGlyThrProCysArgValSerGlyTgPglyThr 141
QY 469 ACCAGAGCCGAGTGTGACCTTCCCTCTGACCTCATGTGGTGGAGTGTCAAGCTCATC 528
DB 142 IleAlaSerProIleAlaArgTyrProThrAlaLeuGlnCysValAsnValAsnIleMet 161
QY 529 TCCCCCGAGGACTGCAGCAAGGTTTACAAAGGACTTACTTGGAAATTCATGCTGCGCT 588
DB 162 SerGluGlnAlaCysHisArgAlaTyrProGlyIleIleThrSerGlyMetValCysAla 181
QY 589 GCATCCCGGACTCCAGAAAGAACCCCTGCAATGCTGAGTCAAGGGGACCGTGTGTGTC 648
DB 182 GlyValProGluGlyGlyLysAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 201
QY 649 AGAGTACCTCCAGAGTCTGTGTCTGGGGAACTTTCCTTGGCGCAACCCCAATGAC 708
DB 202 GlyGlyGlnLeuGlnGlyLeuValSerTrpGlyMetGluArgCysAlaMetProGlyTyr 221
QY 709 CCAGGAGTCTACACTCAAGTGTGCAAGTTTACCAAGTGTGATAAATGACACCATGAAAG 768
DB 222 ProGlyValTyrAlaAsnLeuCysAsnTyrHisSerTrpIleGlnArgThrMetGlnSer 241
QY 769 CAT 771
DB 242 Asn 242

RESULT 14
KLKE HUMAN STANDARD; PRT; 251 AA.
AC Q9POG3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6)
DE (KLK-L6).
GN Name=KLK14; Synonyms=KLK16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

[1]
RN SEQUENCE FROM N.A.
RP Yousef G.M., Diamandis E.P.;
RT "Molecular characterization, mapping, and tissue expression of KLK16,
a hormonally regulated kallikrein-like gene.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP MEDLINE=21250997; PubMed=11352573; DOI=10.1006/geno.2000.6490;
RX Hooper J.D., Bui L.T., Rae F.K., Harvey T.J., Myers S.A.,
RA Ashworth L.K., Clements J.A.;
RT "Identification and characterization of KLK14, a novel kallikrein
serine protease gene located on human chromosome 19q13.4 and expressed
in prostate and skeletal muscle.";
RL Genomics 73:117-122(2001).
[3]
RN SEQUENCE FROM N.A.
RP PubMed=15057824; DOI=10.1038/nature02399;
RX Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
RA Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
RA Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
RA Caenepeel S., Carrano A.V., Caoile C., Chan Y.M., Christensen M.,
RA Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Dettler J.C.,
RA Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M., T.,
RA Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T.,
RA Haydu L., Ho I., Huang W., Idrani S., Jett J., Kadner K., Kimball H.,
RA Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
RA Malfatti S., Martinez D., McCreedy P.M., Medina C., Morgan J.,
RA Nelson K., Nolan M., Orvachenko I., Pitluck S., Pollard M.,
RA Popkie A.P., Fredki P., Quan G., Ramirez L., Rash S., Retterer J.,
RA Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,
RA Slazak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
RA Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,
RA Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,
RA Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
RA Rubin E.M., Lucas S.M.;
RT "The DNA sequence and biology of human chromosome 19.";
RL Nature 428:529-535(2004).
[4]
RN TISSUE SPECIFICITY.
RP MEDLINE=20545474; PubMed=10969073; DOI=10.1074/jbc.M004525200;
RX Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,
RA Clements J.A.;
RT "Tissue-specific expression patterns and fine mapping of the human
kallikrein (KLK) locus on proximal 19q13.4.";
RL J. Biol. Chem. 275:37397-37406(2000).
CC -I- SUBCELLULAR LOCATION: Secreted (Probable).
CC -I- TISSUE SPECIFICITY: High expression in brain, bone marrow and
fetal liver. Also expressed in liver, pancreas, fetal spleen,
prostate and skeletal muscle.
CC -I- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF161221; AAD50773.2; -
CC EMBL; AF283669; AAK48523.1; -
CC EMBL; AF283670; AAK48524.1; -
CC EMBL; AC011473; AAG23260.1; -
CC HSPF; P00760; IEZX.
CC MEROPS; S01.029; -.
CC Genew; HGNC:6362; KLK14.
CC MIM; 606135; -.
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
CC InterPro; IPR009003; Pept_Ser_Cys
```



```
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 24 Activation peptide (Potential).
FT CHAIN 25 251 Kallikrein 14.
FT ACT_SITE 67 67 Charge relay system (By similarity).
FT ACT_SITE 111 111 Charge relay system (By similarity).
FT ACT_SITE 204 204 Charge relay system (By similarity).
FT DISULFID 31 164 By similarity.
FT DISULFID 52 68 By similarity.
FT DISULFID 143 210 By similarity.
FT DISULFID 175 189 By similarity.
FT DISULFID 200 225 By similarity.
SQ SEQUENCE 251 AA; 27452 MW; 9087953BAFA7ED25 CRC64;

Alignment Scores:
Pred. No.: 5,97e-42 Length: 251
Score: 583.00 Matches: 110
Percent Similarity: 61.20% Conservative: 43
Best Local Similarity: 44.00% Mismatches: 89
Query Match: 32.75% Indels: 8
DB: 1 Gaps: 4

US-09-905-083a-30 (1-969) x KIKE_HUMAN (1-251)
QY 28 CTTCTCTGCGCCCTCAGATCTTACTGATCTTACTGCTTCTAGCTTGGAACTCGCAGGAGAA 87
DB 4 LeuLeuThrAlaLeuGlnValLeuAlaIleAlaMet-----ThrGlnSerGlnGlu 20
QY 88 GCCCAGGGTGACAAGATTATTGATCGCGCCCTTGCAGAGGCTGCCACCCATGGCGAG 147
DB 21 AspGlu-----AsnLysIleIleGlyHisThrCysThrArgSerGlnProIlePrgln 39
QY 148 GTGGCCCTGCTCAGTGGC-----AATCAGCTCCACTCGGAGGGCGTCTGTGTCATCAG 201
DB 40 AlaAlaLeuAlaGlyProArgArgPheLeuGlyGlyAlaLeuLeuSerGly 59
QY 202 CGCTGGGTGCTCAGTGGCGCCCTCAGATGATGATGATGATGATGATGATGATGATGATG 261
DB 60 GlnTrpValIleThrAlaAlaHisCysGlyArgProIleLeuGlnValAlaLeuGlyLys 79
QY 262 GATACGCTG-----GGCGACAGGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315
DB 80 HisAsnLeuArgArgTrpGluAlaThrGlnGlnValLeuArgValAlaArgGlnValThr 99
QY 316 CACCCCGGCTACTCCACACAGACCCATGTTATGACCTCATGCTCGTGAAGCTCAATAGC 375
DB 100 HisProAsnTyraAsnSerArgThrHisAspAsnAspLeuMetLeuLeuGlnGlnGln 119
QY 376 CAGGCGCAGCTGCTATCCATGATGTAAGAAAGTCAAGCTGCCCTCCGCTGGCAACCCCT 435
DB 120 ProAlaArgIleGlyArgAlaValArgProIleGluValThrGlnAlaCysAlaSerPro 139
QY 436 GGAACCACTGACTGCTCCGCTGGCGCTGATGATGATGATGATGATGATGATGATGATGAT 495
DB 140 GlyThrSerCysArgValSerGlyTrpGlyThrIleSerSerProIleAlaArgTrpPro 159
QY 496 TCTGACCTCATGCTGGTGTGATGCTCAAGTCTCATCTCCCGCAGGACTCGCAGAGGTTTAC 555
DB 160 AlaSerLeuGlnCysValAsnIleAsnIleSerProAspGluValCysGlnLysAlaTy 179
QY 556 AAGGACTTACTGGAAATTCATGCTGTGCGCTGCGATCCCGATCCCAAGAAACGCC 615
DB 180 ProArgThrIleThrProGlyMetValCysAlaGlyValProGlnGlyGlyLysAspSer 199
QY 616 TCCATGGTGACTCAGGGGCGCTGTGTGTGTCAGAGGTACCTCCAGGTCTGTGTGCC 675

Db 200 CysGlnGlyAspSerGlyGlyProLeuValCysArgGlyGlnLeuGlnGlyLeuValSer 219
QY 676 TGGGGAACCTTCCCTTGGCGCCCAACCCCAATGACCCAGAGTCTACCTCAAGTGTGCAAG 735
DB 220 TrpGlyMetGluArgCysAlaLeuProGlyTyrProGlyValTyrThrAsnLeuCysLys 239
QY 736 TTCACCAAGTGGATAAATGACACCATGAAA 765
DB 240 TyrArgSerTrpIleGluGluThrMetArg 249

RESULT 15
Q6B089 PRELIMINARY; PRT; 251 AA.
AC Q6B089;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kallikrein 14, preproprotein.
GN Name=KLK14;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; BC074905; AAH74905.1; -.
DR EMBL; BC074904; AAH74904.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; TRYPSIN_DOM; 1.
DR InterPro; IPR00135; TRYPSIN_SER; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 251 AA; 27507 MW; C353A7D8B1960BD2 CRC64;

Alignment Scores:
Pred. No.: 5,97e-42 Length: 251
```

Score: 583.00 Matches: 110  
Percent Similarity: 61.20% Conservative: 43  
Best Local Similarity: 44.00% Mismatches: 89  
Query Match: 32.75% Indels: 8  
DB: 2 Gaps: 4

US-09-905-083A-30 (1-969) x Q6B089 (1-251)

QY 28 CTTCTCCTCCCTCCCTCAGATCTTACTGCTATCTTATGCTTGGAAACTGCAGGAGAA 87  
Db 4 LeuLeuThrAlaLeuGlnValLeuAlaIleMet-----ThrArgSerGlnGlu 20  
QY 88 GCCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCACCCATGGCAG 147  
Db 21 AspGlu--AsnLysIleIleGlyTyrThrCysThrArgSerSerGlnProIleGln 39  
QY 148 GTGGCCCTGCTCAGTGGC-----AATCAGCTCCACTGGGAGGGCTCTGTCTCAATGAG 201  
Db 40 AlaAlaLeuLeuAlaGlyProArgArgPheLeuCysGlyGlyAlaLeuLeuSerGly 59  
QY 202 CGCTGGGTGCTCACTGCGCCCTCCACTGCAAGATGATGATACACCGTGCACCTGGGCAGT 261  
Db 60 GlnTrpValIleThrAlaAlaHisCysGlyArgProIleLeuGlnValAlaLeuGlyLys 79  
QY 262 GATACGCTG-----GGCGACAGGAGAGCTCAGAGGATCAAGGCTCGAAGTCATTCGCG 315  
Db 80 HisAsnLeuArgArgTrpGluAlaThrGlnGlnValLeuArgValValArgGlnValThr 99  
QY 316 CACCGCGGCTACTCCACAGACCCATGTTAATGACCTCATGCTCTGTAAGCTCAATAGC 375  
Db 100 HisProAsnTyrAsnSerArgThrHisAspAsnAspLeuMetLeuLeuGlnGlnGln 119  
QY 376 CAGGCGAGGCTGCTCATCTCCATGCTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCT 435  
Db 120 ProAlaArgIleGlyArgAlaValArgProIleGluValThrGlnAlaCysAlaSerPro 139  
QY 436 GGAACACCTGTACTGTCTCGGCTGGGGGCACTACACAGAGCCAGATGTGACCTTTCCC 495  
Db 140 GlyThrSerCysArgValSerGlyTrpGlyThrIleSerSerProIleAlaArgTyrPro 159  
QY 496 TCTGACCTCATGTGGTGGATGCTCAAGCTCATCTCTCCCCCAGGACTGCACGAGGTTTAC 555  
Db 160 AlaSerLeuGlnCysValAsnIleAsnIleSerProAspGluValCysGlnLysAlaTyr 179  
QY 556 AAGGACTTACTGAAAATTCCATGTGTGCTGGCTGCCATCCCGACTCCCAAGAAAAACGCC 615  
Db 180 ProArgThrIleThrProGlyMetValCysAlaGlyValProGlnGlyGlyAspSer 199  
QY 616 TGCATGTGTACTCAGGGGACCGTTGGTGTGACAGAGTACCTTCGAAGGTCTGTGTCC 675  
Db 200 CysGlnGlyAspSerGlyProLeuValCysArgGlyGlnLeuGlnGlyLeuValSer 219  
QY 676 TGGGGAACCTTCCCTGGCGCCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735  
Db 220 TrpGlyMetGluArgCysAlaLeuProGlyTyrProGlyValTyrThrAsnLeuCysLys 239  
QY 736 TTCACCAAGTGGATAAATGACACCATGAAA 765  
Db 240 TyrArgSerTrpIleGluGluThrMetArg 249

Search completed: July 14, 2005, 17:31:58  
Job time : 211.5 secs